

AD _____

(Leave blank)

Award Number: W81XWH-07-1-0306

TITLE: PROSPECT: Profiling of Resistance Patterns & Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax and Therapeutic Target Identification

PRINCIPAL INVESTIGATOR:

Waun Ki Hong, M.D.

CONTRACTING ORGANIZATION:

The University of Texas M. D. Anderson Cancer Center
Houston, Texas 77030

REPORT DATE: June 2010

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

☒ Approved for public release; distribution unlimited

☐ Distribution limited to U.S. Government agencies only;
report contains proprietary information

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 01-06-2010		2. REPORT TYPE Annual		3. DATES COVERED (From - To) 1 June 2009 - 31 May 2010	
4. TITLE AND SUBTITLE PROSPECT (Profiling of Resistance Patterns & Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax and Therapeutic Target Identification)			5a. CONTRACT NUMBER W81XWH-07-1-0306		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Waun Ki Hong, M.D. whong@mdanderson.org			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The University of Texas M.D. Anderson Cancer Center 1515 Holcombe Blvd., Unit 432 Houston, TX 77030			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research And Materiel Command Fort Detrick, MD 21702-5012			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACTWe will develop a high throughput therapeutic-target focused (TTF) profiling platform and will this with tumor genome wide mRNA profiling and with serum or plasma profiling of phosphopeptides and DNA. We will use these molecular profiles to help define how various molecular factors alone and in combination relate to resistance to therapy, to prognosis, and to metastatic patterns at relapse. Using tumor and blood samples from non-small cell lung cancer (NSCLC) patients as well as NSCLC cell lines with defined chemotherapy resistance patterns, we will examine how molecular profiles may confer resistance and will identify new, potential therapeutic targets. The PROSPECT approach will be novel in that we will assess tumors from NSCLC patients undergoing surgical resection after having received neoadjuvant therapy as a model of resistance. Tumor surviving neoadjuvant therapy would be expected to be enriched for resistant cells. We will define what combinations of targeted therapies are most effective against resistant cell lines with similar molecular profiles, and this will drive later clinical trials (beyond the scope of this Program). Similar studies will be conducted in patients with mesotheliomas undergoing surgical resection of tumor after neoadjuvant therapy with the new Src inhibitor dasatinib.					
15. SUBJECT TERMS Lung cancer, mesothelioma, target-focused profiling, resistance					
16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 378	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code)

TABLE OF CONTENTS

INTRODUCTION	4
PROGRESS REPORT (BODY)	4
<i>Project 1</i>	4
<i>Project 2</i>	11
<i>Project 3</i>	16
<i>Project 4</i>	30
<i>Project 5</i>	33
<i>Pathology Core</i>	38
<i>Biostatistics Core</i>	42
KEY RESEARCH ACCOMPLISHMENTS	47
REPORTABLE OUTCOMES	50
CONCLUSIONS	53
REFERENCES	55
APPENDICES	56

Appendix A - Abstracts and Publications

Appendix B - Biostatistics Database Screenshots

INTRODUCTION

Lung cancer is the leading cause of cancer death in the world. Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancer cases. Only 15% of patients diagnosed with lung cancer survive five years from diagnosis. Therapy for advanced disease increases average life expectancy by only a few months, and slightly improves quality of life. Similarly, adjuvant chemotherapy for resected disease has only a modest impact on survival rates. More effective therapy is needed. We believe that applying state-of-the-art molecular tools to carefully conducted clinical trials will lead to the identification of molecular mechanisms that contribute to lung cancer therapeutic resistance and that drive prognosis, and that this in turn will lead to the development of drugs with novel biological and therapeutic functions. Therefore, we have undertaken a translational research program named **PROSPECT: Profiling of Resistance Patterns & Oncogenic Signaling Pathways in Evaluation of Cancers of the Thorax and Therapeutic Target Identification**. The goal of PROSPECT is to use therapeutic target-focused (TTF) profiling along with genome-wide mRNA and serum phosphopeptide profiling to identify and evaluate molecular targets and pathways that contribute to therapeutic sensitivity or resistance, prognosis, and recurrence patterns, and to use this information to guide formulation of new rational therapeutic strategies for NSCLC and mesotheliomas. In the Program, we have 5 research projects and 3 Cores to address 3 central issues: therapeutic resistance, prognosis and new therapeutic targets and strategies.

PROGRESS REPORT (BODY):

Project 1: Therapeutic target-focused (TTF) profiling for the identification of molecular targets and pathways that contribute to drug sensitivity or resistance *in vitro* and the development of rational treatment strategies for NSCLC.

(Leader: Dr. John Heymach; Co-Leader: Dr. John Minna)

Hypotheses:

We hypothesize that a broad, systematic molecular profiling of NSCLC cell lines, using both TTF and global approaches, will lead to the following results:

1. The identification of new potential therapeutic targets for NSCLC
2. The development of predictive markers for *in vitro* sensitivity to targeted agents, which will form the starting point for the development of a predictive model of *in vivo* sensitivity using clinical specimens as described in Aim 3.
3. Insights into the molecular mechanism underlying therapeutic resistance and into the relationship of resistance mechanisms to factors innately affecting tumor growth rate and prognosis
4. Identification of readily translatable therapeutic strategies to combat these resistance mechanisms.

Specific Aims:

In this project, we will develop and validate a novel therapeutic target-focused (TTF) profiling platform at M.D. Anderson Cancer Center. The platform will provide a high throughput, quantitative, scalable, and highly sensitive set of assays to assess activation of key signaling pathways (e.g., PI3K/AKT, STAT, RAS-RAF-ERK) as well as other potential therapeutic targets such as receptor tyrosine kinases (RTKs). It will be coupled with global profiling of gene expression using Affymetrix 2.0 array. These molecular profiles will then be coupled with information from a broad drug and therapeutic target siRNA (DATS) screen to develop markers for predicting drug sensitivity *in vitro* based on molecular profiles, elucidate the molecular determinants of sensitivity or resistance to a given therapeutic agent, and identify potential

therapeutic targets for tumor cells resistant to a given agent. This project lays the foundation for Project 3, where the same TTF and global profiling approaches will be used to characterize clinical tumor specimens and investigate molecular markers identified in this project, for Project 4, in which the profiles and therapeutic targets for mesothelioma will be explored, and for Project 2, in which the profiles will be correlated with patient prognosis and metastatic patterns. The specific aims of this project are as follows:

Specific Aim 1: To develop a TTF profile for assessing critical signaling pathways and potential therapeutic targets, and to apply TTF and gene expression profiling to NSCLC and mesothelioma cell lines.

- 1.1. Development and technical validation of a TTF profile using reverse phase lysate arrays (RPPA) and multiplexed bead array technology.
- 1.2. Application of TTF profiling to a cell line panel representing malignant (NSCLC and mesothelioma) and non-malignant (endothelial and stromal cells, normal bronchial epithelium) cell types.
- 1.3. Gene expression profiling of the cell line panel using Affymetrix microarrays.
- 1.4. Correlation of TTF and gene expression profiles from the cell line panel to determine gene expression signatures that correlate with activation of individual proteins (e.g., EGFR activation) and critical signaling pathways (e.g., RAS pathway activation).

Specific Aim 2: To determine the sensitivity of the cell line panel to the selected drug and therapeutic target siRNA (DATS) screen.

- 2.1. Screening of the cell line panel for sensitivity to a panel of 20-25 targeted agents and standard chemotherapy agents.
- 2.2. Screening of the cell line panel using siRNA representing potential therapeutic targets, including molecules targeted by specific agents in Aim 2.1 (e.g., EGFR, IGFR-1, etc.) and potential therapeutic targets for which drugs are not currently available (e.g., RTKs for which drugs are currently in development).
- 2.3. Comparison of *in vitro* and *in vivo* profiles (TTF and global) and drug sensitivity in selected NSCLC cell lines and xenografts grown from the same lines.

Specific Aim 3: Development of markers for predicting drug and targeted siRNA sensitivity *in vitro* based on TTF and molecular profiles, and identification of candidate therapeutic targets in chemotherapy-resistant lines.

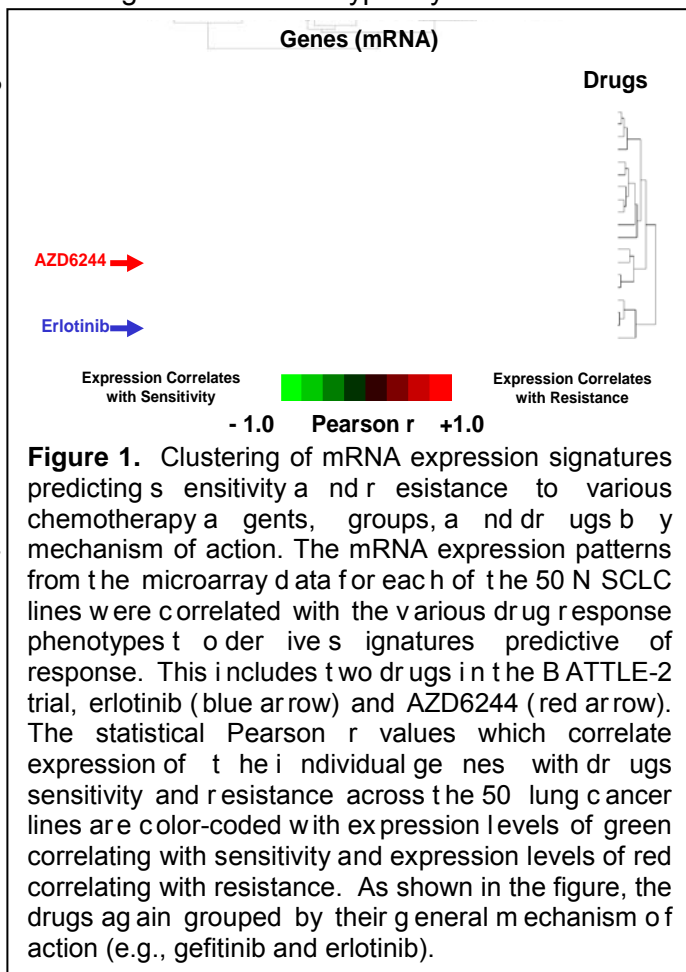
Summary of Research Findings

We have established a cell line bank with a standardized set of 75 NSCLC, 30 SCLC, and 5 immortalized human bronchial epithelial cell (HBECS) lines that represent a range of histological tumor types including adenocarcinoma, squamous, and bronchioalveolar. These cell lines have been DNA-fingerprinted and mycoplasma- and virus-tested to insure they are of the correct origin and are contamination-free. Further, we developed over 70 different NSCLC xenografts, including 30 derived from human NSCLC cell lines, stably expressing luciferase to allow for bioluminescence imaging (BLI). We generated more than 40 heterotransplant models, with similar histology to the primary tumor counterparts, representing major histological NSCLC subtypes (in collaboration with Drs. Wistuba and Mao). We have completed genome-wide mRNA expression profiling using Affymetrix HGU133(A, B, or Plus2) or Illumina WG6-v2 gene chips and protein profiling by reverse-phase proteomic arrays (RPPA) on all of these cell lines; expression data was correlated with *in vitro* drug response (by MTS and colony formation assays) to a number of chemotherapies and targeted agents (**Figure 1**).

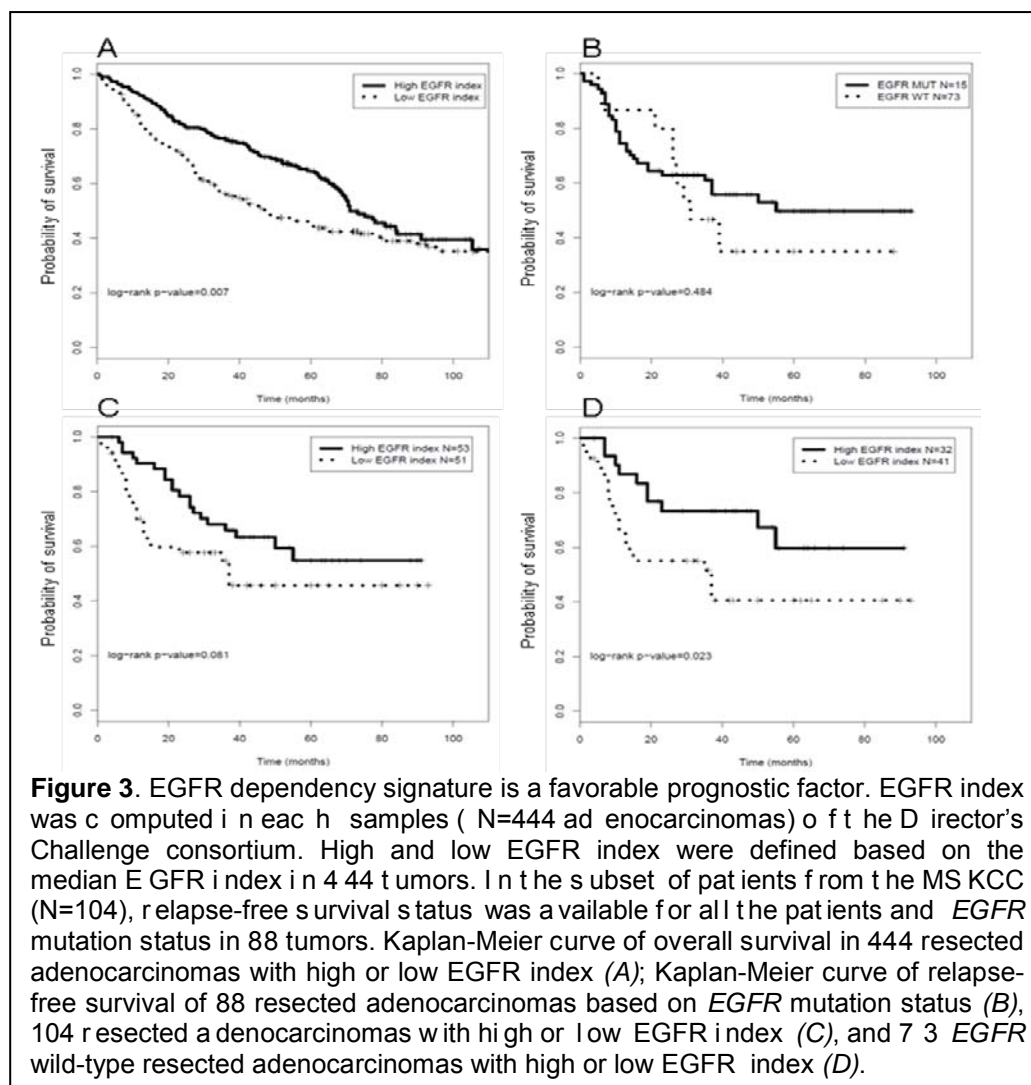
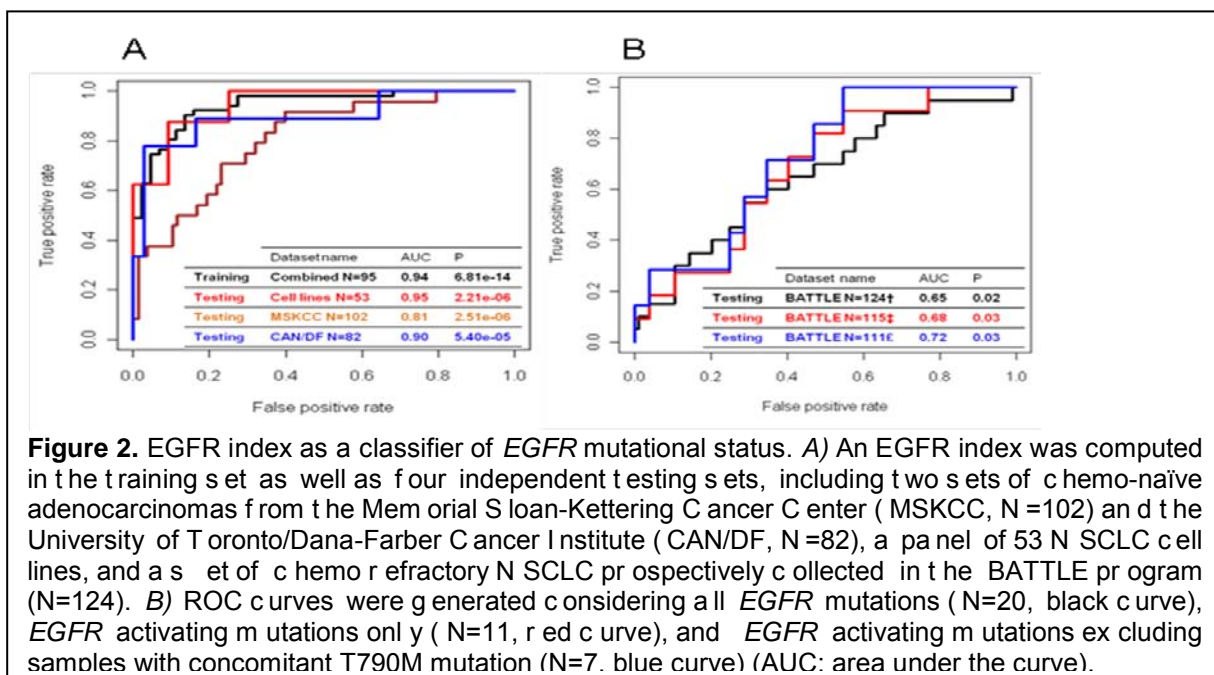
We identified signatures that predict response to these therapeutics and those results are reported below. These results will be validated in xenografts and in clinical samples from patients treated with these drugs.

An *EGFR* mutation signature is prognostic in *EGFR* wild-type lung adenocarcinomas and identifies *Metastasis associated in colon cancer 1 (MACC1)* as an *EGFR* mutant-associated regulator of MET. NSCLC bearing *EGFR* activating mutations are typically sensitive to *EGFR*

tyrosine kinase inhibitors (TKIs) and have increased MET levels through multiple mechanisms. We developed an *EGFR* mutation gene expression signature to identify genes associated with the *EGFR*-dependent phenotype. Gene expression and *EGFR* status were measured in 95 adenocarcinomas (training set), 53 NSCLC cell lines (testing set 1), two sets of chemo-naïve adenocarcinomas (82 and 102), lung adenocarcinomas (testing set 2 and 3), and one set of chemorefractory NSCLC collected prospectively in the BATTLE program (N=139, testing set 4). An *EGFR* index was computed in each sample, using log₂-based expression values, and was defined as the average of log₂-based expression values of the genes up-regulated in *EGFR*-mutant samples minus the average of the log₂-based expression values of the genes down-regulated in *EGFR* wild-type samples, as defined by SAM algorithm in the combined training set. Receiver operating characteristic curves were used to evaluate the value of *EGFR* index as a classifier of *EGFR* mutational status. Proteomic profiling analysis, using 170 antibodies, was generated by RPPA in 74 NSCLC cell lines. Immunostaining for MACC1 and MET of a tissue-microarray, including 287 NSCLC, were studied.



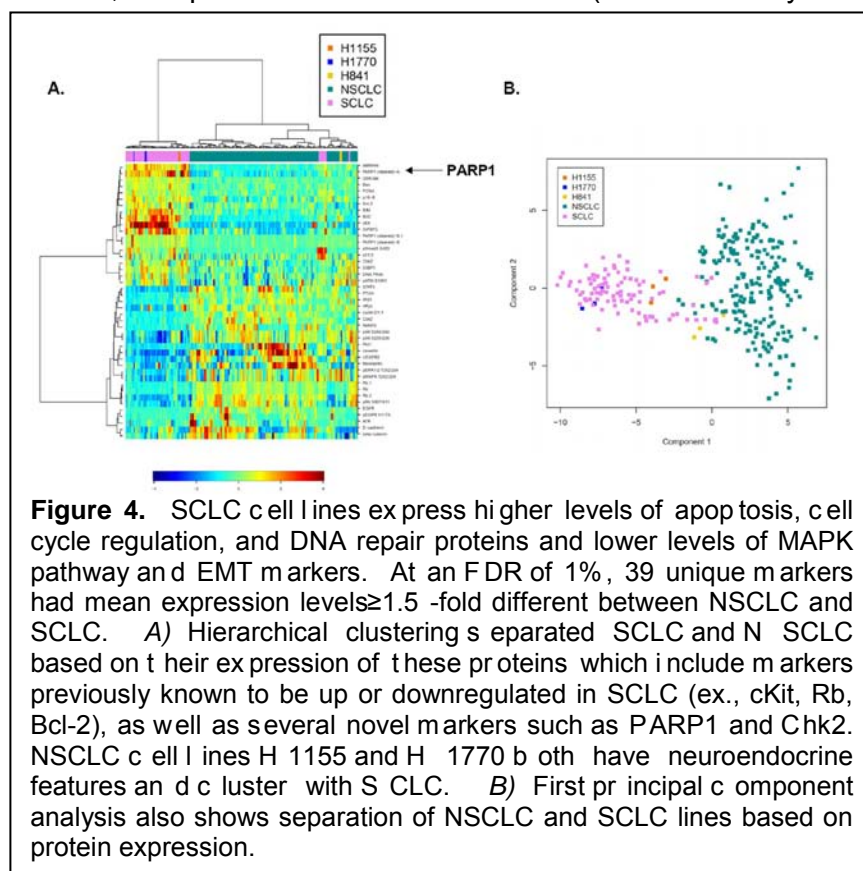
Using SAM algorithm, we identified a set of 45 genes significantly associated with *EGFR* status in the training set, with a fold-change set at 2. The gene-set was used to compute a signature that was shown to be predictive of *EGFR* status in 4 testing sets, with a strong association in cell lines and chemo-naïve adenocarcinomas (**Figure 2**). Pathway analysis of the genes showed that *EGFR* mutations were marked by increased endocytosis and decreased mitosis genes. The signature correlated with *EGFR* index was significantly correlated with sensitivity to both erlotinib ($r=-0.44$, Pearson's product-moment correlation p -value= $1.2e-03$) and gefitinib ($r=-0.58$, Pearson's product-moment correlation p -value= $4.2e-05$). Gene expression and clinical annotations from the Directors' Challenge Consortium were downloaded from the NCI website (N=444 adenocarcinomas). The *EGFR*-dependent signature was significantly associated with survival in early stage adenocarcinomas, even in the subgroup with *EGFR* wild-type tumors (**Figure 3**).



Our current study identified *MACC1* as the most significantly up-regulated gene in *EGFR* mutant. *MACC1* and *MET* expression levels, *in vitro* and *in vivo*, were highly correlated both at the mRNA and protein levels. A correlation heatmap of the proteomic profiling showed a strong correlation of *MACC1* with *MET*, downstream signal transducers (STATs), markers of the epithelial phenotype, a classical feature of *EGFR* mutant tumors, and *RAB25*, a member of the RAS superfamily of small GTPases that are involved in endocytosis.

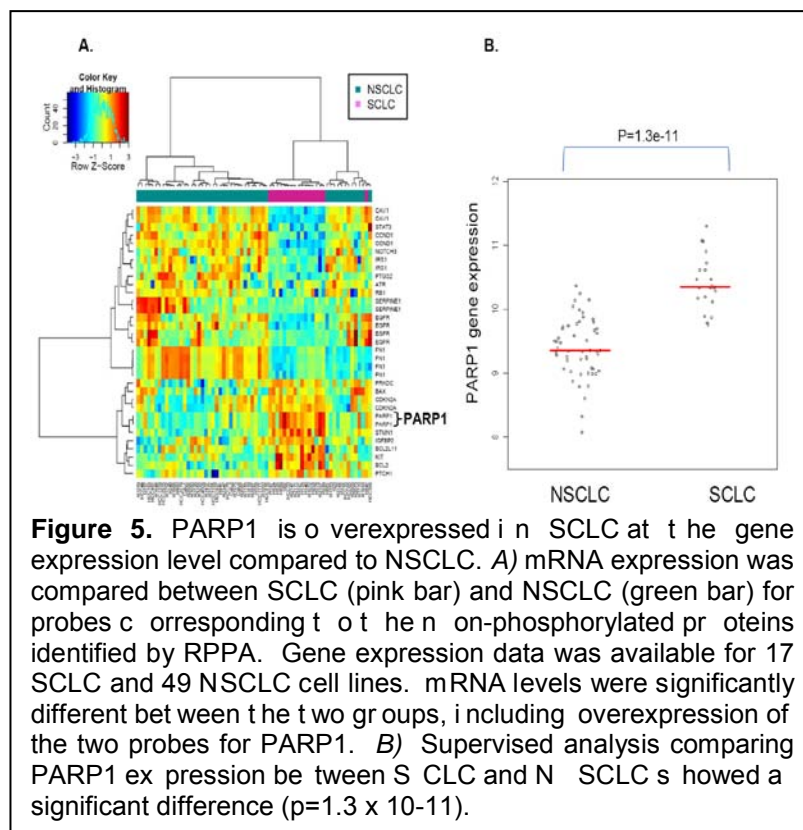
Finally, we showed that *MACC1* knockdown *in vitro* reduced *MET* levels, using both siRNA and shRNA experiments. A gene expression signature associated with *EGFR*-dependent phenotype was developed. The signature was shown to be prognostic regardless of *EGFR* status. The results also suggest *MACC1* to be a regulator of *MET* in NSCLC.

Protein expression profiling identifies high PARP1 expression in SCLC. Protein lysates from 34 SCLC and 74 NSCLC cell lines were harvested after 24h growth in full serum (10%), serum starved (0%), and serum stimulated conditions and printed in serial dilution series on the RPPA slides. Protein levels of 176 proteins and phosphoproteins in key signaling pathways commonly dysregulated in cancer were identified by RPPA analysis. Using a mixed linear model, we identified 17 proteins whose mean expression levels in SCLC were ≥ 1.5 fold higher than in NSCLC, independent of media conditions (false discovery rate $< 1\%$, corresponding p-value < 0.001) (**Figure 4**).



Proteins upregulated in SCLC included those previously shown to be overexpressed such as cKit and Bcl2. However, a majority of proteins that we identified were not previously well characterized in SCLC. Among those proteins overexpressed in SCLC, we identified a significant elevation in several DNA repair proteins and cell cycle regulators such as PARP1, XRCC1, ATM, Chk2, and 53BP1. Among these, PARP1 was markedly increased, with 4-fold higher median expression in SCLC cell lines, as compared to NSCLC. The mRNA expression levels of non-phosphorylated proteins

differentially expressed between SCLC (n=17) and NSCLC (n=49) were then compared in the cell lines using Affymetrix U133A, B, and Plus 2.0 arrays. As with the protein data, PARP1 mRNA levels were significantly higher in SCLC. (**Figure 5A**). Analysis of an independent set of SCLC cell lines also demonstrated markedly higher expression of PARP1 in SCLC as compared to NSCLC (**Figure 5B**). In addition, we compared levels of PARP1 in publicly available gene expression data from patient tumors with either SCLC or NSCLC. Although the numbers of



SCLC tumors are small, there was a significantly higher level of PARP1 expression in these tumors as compared to NSCLC (gene expression data from Bhattacharjee et al.).

Based on this data and promising clinical results from PARP inhibitors in clinical trials of other solid tumors, we are interested in PARP1 as a potential new therapeutic target in SCLC. In addition, preclinical data has suggested that drug inhibition of PARP1 may also be synergistic with other DNA-damaging agents such as cisplatin and radiation, which would make them particularly promising for use in SCLC. Because of this, we have preliminarily tested a small number of SCLC with high PARP1 expression and

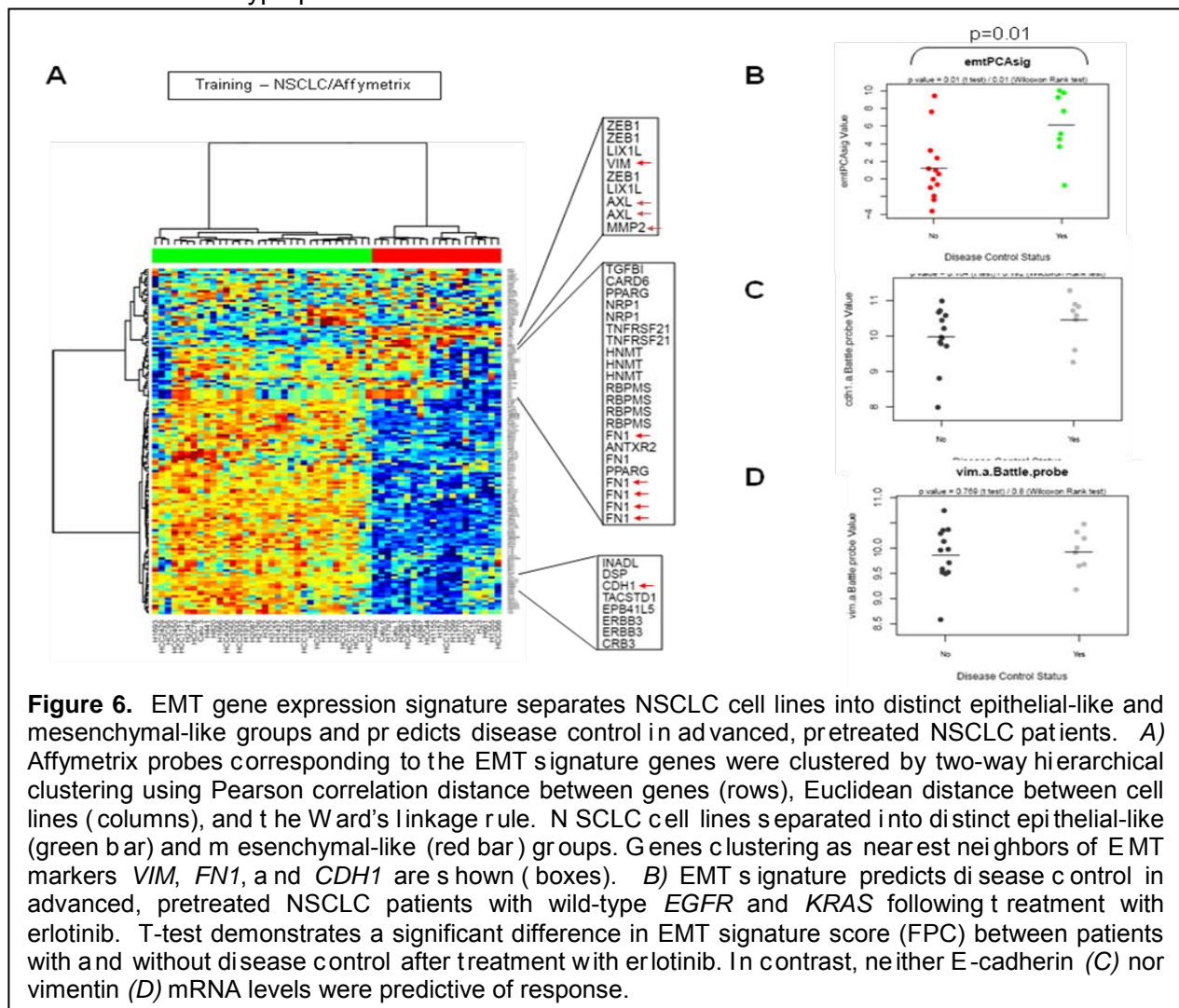
compared them to NSCLC cell lines with low PARP1 expression. We also tested one SCLC cell line (H841), which had low levels of PARP and clustered with NSCLC cell lines based on protein and gene expression profiles in our previous profiling studies. We found that SCLC cell lines with high PARP1 expression appear to be sensitive to PARP1 inhibition, and that their degree of response is correlated with their relative level of PARP1 expression (high or intermediate). In contrast, NSCLC cell lines and H841 that have low PARP1 expression were relatively resistant to PARP inhibition. This data suggest that PARP1 may be a clinically active target in SCLC, and we plan to investigate these findings further in preclinical and clinical samples from patients with SCLC.

An epithelial to mesenchymal transition (EMT) gene expression signature identifies Axl as an EMT marker in NSCLC head and neck cancer (HNC) lines and predicts response to erlotinib. EMT is associated with loss of cell adhesion molecules, such as E-cadherin, and increased invasion, migration, and proliferation in epithelial cancers. In NSCLC, EMT is associated with worse prognosis and resistance to EGFR inhibitors. Despite these clinical implications, no gold standard exists for classifying a cancer as either epithelial or mesenchymal in nature. The goal of this study was to develop a robust EMT gene expression signature and test its correlation with drug response.

The EMT signature was derived in 54 DNA-fingerprinted NSCLC cell lines profiled on Affymetrix U133A, B, and Plus2.0 arrays and tested on the Illumina WGv2 and WGv3 platforms, and in an independent set of head and neck cancer lines. E-cadherin and other protein levels were quantified by RPPA and correlated with the first principal component of the EMT signature. IC50s were determined for NSCLC cell lines by MTS assay.

Expression of 76 genes comprising the EMT signature correlated with mRNA expression of known EMT markers E-cadherin, vimentin, N-cadherin, or fibronectin 1, and was bimodally distributed across the NSCLC panel. Classification of the NSCLC lines as epithelial or mesenchymal by the EMT signature was verified for 51/52 cell lines tested on both Affymetrix and Illumina platforms (**Figure 6**). In an independent validation set of 62 HNC lines, the signature identified a subset of six mesenchymal cell lines. The EMT signature correlated well with E-cadherin protein levels in NSCLC ($r=0.90$) and HNC ($r=0.73$). mRNA levels for Axl, a tyrosine kinase receptor associated with EMT in breast cancer, had the most negative correlation with E-cadherin ($r=-0.45$) of any signature gene after ZEB1 and vimentin, and was positively correlated with vimentin ($r=0.60$) and N-cadherin ($r=0.54$) expression. Higher Axl total protein was confirmed in NSCLC and HNC mesenchymal-like cell lines. Mesenchymal phenotype (classified by the EMT signature) was more strongly correlated with NSCLC erlotinib resistance ($p=0.028$) than E-cadherin mRNA or protein level.

In conclusion, an EMT gene expression signature accurately classifies cell lines as epithelial- or mesenchymal-like across three microarray platforms and two cancer types, and identifies Axl as a novel EMT marker in NSCLC and HNC. The EMT signature was a better predictor of erlotinib resistance than single mRNA or protein markers, such as E-cadherin, both in cell lines and in *EGFR/KRAS* wild-type patients.



Key Research Accomplishments

- Completed protein profiling and gene expression profiling for 75 NSCLC, 30 SCLC, and 5 HBEC cell lines.
- Developed an *EGFR* mutation signature that is prognostic in *EGFR* wild-type lung cancer.
- Identified M *ACC1* as highly expressed in lung cancer as an *EGFR* mutant-associated regulator of MET.
- Identified PARP as being highly expressed in SCLC and as a clinically relevant target.
- Developed an EMT gene expression signature that identifies epithelial and mesenchymal groups in NSCLC and HNC cell lines
- Identified a novel EMT marker, Axl, in NSCLC and HNC.

Conclusions

Protein and gene expression profiling of a large number of cell lines has allowed us to identify a prognostic *EGFR* dependent signature as well as key signaling molecules in NSCLC (*MACC1*) and SCLC (PARP) that were highly upregulated in those tumor types. Further characterization of these proteins will identify their role in the development of cancer. Also, as in the case with PARP, investigation into the inhibition of these targets may prove to be an effective strategy for the treatment of these diseases. We developed an EMT gene expression signature that classified cell lines into epithelial- or mesenchymal-like group in NSCLC and HNC. This signature was able to predict response to erlotinib in cell lines and *EGFR/KRAS* wild-type patients. Additionally, a novel EMT marker, Axl, was identified through this analysis and may play an important role in the transformation of these tumors as well serve as a potential therapeutic target. These profiles will allow for multiple biomarker analysis and the identification of intracellular signaling pathways that contribute to the sensitivity or resistance to therapeutics. Our findings will be further validated by correlating these gene expression and proteomic profiles of tumor samples with clinical outcomes in samples from the BATTLE-1 trial. Development of predictive markers will assist in guiding treatment selection as well as identify new targets in lung cancer.

Project 2: Tumor molecular profiles in patients with operable non-small cell lung cancer (NSCLC): impact on stage, prognosis, and relapse pattern.

(Leaders: Drs. David Stewart, Jack Roth; Co-Leaders: Drs. Roy Herbst, Edward Kim, Katherine Pisters, Stephen Swisher)

Hypotheses:

We hypothesize that:

1. In tumors from patients with NSCLC, patterns of co-expression of molecules that modulate cell proliferation, survival, angiogenesis, invasion, metastasis and apoptosis will substantially influence tumor stage and size at the time of diagnosis, and will largely define patient prognosis.
2. Impact of adjuvant and neoadjuvant therapies on disease-free, progression-free, and overall survival will vary across prognostically distinct groups.
3. Specific molecular signatures in primary tumors will predict both metastatic patterns at relapse and molecular profiles of recurrent tumors, and this could help guide adjuvant strategies and therapeutic strategies at relapse.

Specific Aims:

Aim 1: To define characteristic TTF/gene expression profiles of prognostically distinct subpopulations of patients with resectable NSCLC, and to assess the extent to which these molecular profiles correlate with tumor stage and/or size.

The main goal of this aim is to use 150 archival NSCLC tumor samples from our tissue bank (with corresponding clinical data) and to prospectively collect tumor samples, blood samples, and clinical data from 300 additional patients undergoing surgical resection of NSCLC. The tissue and blood samples will be used by Project 3 and the Pathology Core to generate comprehensive TTF/gene expression molecular profiles using methods developed in Project 1. We will construct Kaplan-Meier estimated survival curves for disease-free survival, progression-free-survival, and overall survival, and will use Cox proportional hazards models and recursive partitioning methods to identify important biomarkers and prognostically distinct subpopulations. We will also correlate TTF/gene expression molecular profiles with initial tumor size and stage. In addition, we will explore the feasibility of using nonlinear regression analyses of semilog plots of % disease-free survival, % progression-free survival, and % overall survival vs time to facilitate identification of prognostically distinct subpopulations with characteristic TTF/gene expression molecular profiles.

Aim 2: To assess the impact of adjuvant and neoadjuvant chemotherapy on disease-free survival, progression-free survival, and overall survival in prognostically distinct subgroups, and to provide tumor, blood and clinical data to Project 3 for an assessment of factors contributing to resistance to chemotherapy and to Project 5 for assessment of profiling of EGFR and related molecules by new quantum dot technologies.

Of the 450 patients included in the project, we will assess 100 new prospectively recruited patients who will receive neoadjuvant therapy, 100 patients who will receive postoperative adjuvant therapy (including approximately 20 tumor bank patients and 80 new patients), and 250 patients who did not receive adjuvant or neoadjuvant therapy (including approximately 130 tumor bank patients and 120 new patients). We will collect patient clinical data on all 450 patients and will collect blood samples on the 300 new, prospectively recruited patients. Tumor and blood samples and clinical data will be provided to Project 3 for studies of therapeutic resistance and to Project 5 for assessment of profiling of epidermal growth factor receptor (EGFR) and related molecules by new quantum dot technologies, while in Project 2 we will assess impact of adjuvant and neoadjuvant therapy on outcome in each prognostic group.

Aim 3: To correlate TTF/gene expression molecular profiles in the primary tumor with metastatic patterns and with tumor molecular profiles at relapse.

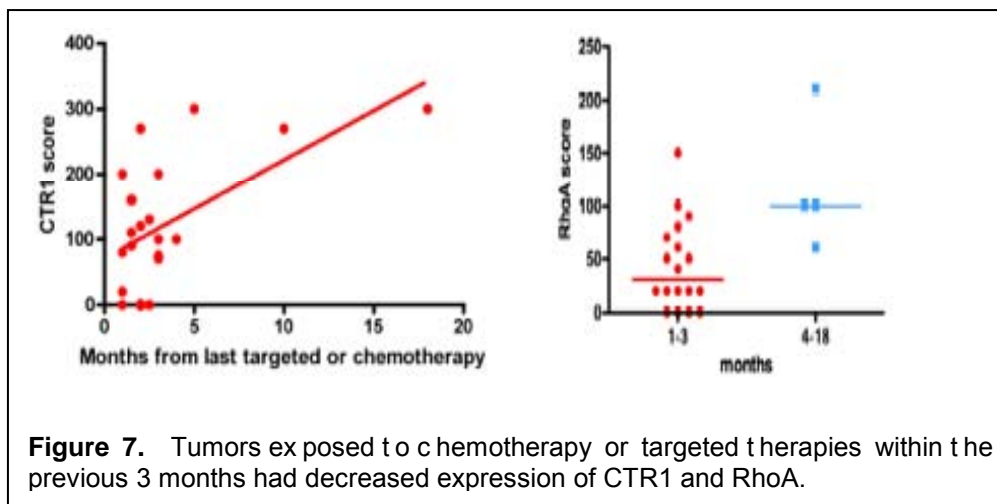
For patients who relapse, we will define metastatic sites at relapse, obtain tumor tissues from selected patients who undergo biopsies to confirm relapse, and define TTF/gene expression molecular profiles in the patients' original primary tumor specimens that predict sites of later relapse (and in particular that predict relapse in brain). We will also assess whether tumor at relapse is enriched for particular molecular characteristics that may promote metastasis when compared to the primary tumor, and will assess the extent to which TTF/gene expression molecular profile at diagnosis may help guide choice of therapies at relapse.

Summary of Research Findings

As outlined in more detail in the reports from Project 3 and the Pathology Core, we have assessed frozen and formalin-fixed, paraffin-embedded (FFPE) tumor specimens from our Tumor Bank for 750 patients who underwent surgery between 1996 and July 2007. DNA and RNA were extracted, and quality was assessed (RIN). From these, 249 cases were selected based on predefined criteria. Of these, 40 tumors corresponded to neoadjuvant-treated tumors. Specifically, the cases selected had a RIN number ≥ 4 and were adenocarcinomas or squamous cell carcinomas of the lung with tumor (vs normal tissue) comprising $\geq 70\%$ of the specimen, and with malignant cells (vs stromal cells) comprising $\geq 30\%$ of the tumor portion. Cases selected have been assessed for microRNA, mRNA, and DNA aCGH profiling, and prepared for proteomics analysis for 170 proteins. Mutation analysis of *KRAS* (codons 12, 13 and 61) and *EGFR* (exons 19 and 21) is underway. To date, from 227 NSCLC cases tested, *KRAS* mutations have been detected in 40 out of 136 (29%) adenocarcinomas and in none of 91 squamous cell carcinomas examined. Additionally, DNA has been prepared for mutation analysis of a 20-gene panel using the MALDI-TOF Mass Spectrometry-based (SNP) analysis, Sequenom®.

Some early analyses have revealed that several of the proteins assessed to date by IHC (e.g., carbonic anhydrase IX, Nrf2, and Keap1) have correlated with prognosis and outcome of patients with surgically resected NSCLC tumors and treated with adjuvant therapy. In a preliminary assessment of gene expression arrays with FFPE tissues, we were able to identify a gene signature that distinguished a good-prognosis group from a poor-prognosis group, and this signature remained significantly associated with prognosis even after correction for other variables in a multivariate analysis ($p=0.025$). In addition, for patients treated with neoadjuvant chemotherapy, OS and RFS were correlated to the amount of tissue that remained viable after the neoadjuvant therapy, demonstrating that patients with the greatest tumor response from chemotherapy had the best survival.

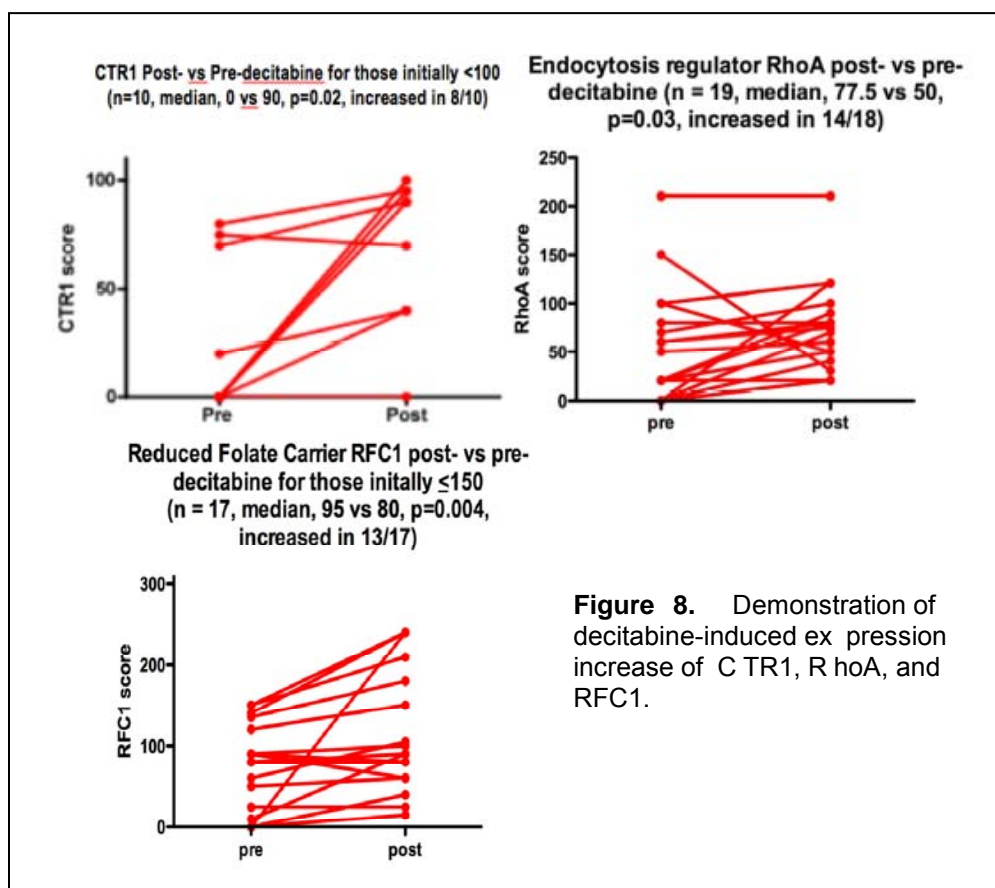
Collection of prospective tumor samples is also going well. We had proposed to collect 300 blood and tissue samples prospectively, and have surpassed that goal. Since the activation of the PROSPECT protocol on August 2007, we have collected fresh tissue specimens from 475 NSCLC surgically resected cases. From those, snap frozen normal and tumor tissue have been collected in all cases. We originally proposed to collect tumor samples prospectively from 100 patients who received neoadjuvant chemotherapy, and to date we have collected 117. Hence, we are ahead of schedule on specimen procurement for the project. In Project 3,



expression profiles in tumors from

patients in Project 2 who did or did not receive neoadjuvant chemotherapy will be compared. Viable tumor tissue obtained after neoadjuvant chemotherapy will be regarded as a model of acquired resistance.

In preliminary work assessing potential resistance factors in a variety of types of human tumor samples, we found that tumors exposed to chemotherapy or targeted therapies within the previous 3 months had decreased expression of the copper/platinum transporter CTR1 [1] and of the endocytosis regulator RhoA[2] (**Figure 7**), suggesting a mechanism by which exposure to a broad range of agents could secondarily lead to resistance to cisplatin and carboplatin. We also demonstrated that the DNA demethylating agent decitabine could increase expression of each of CTR1, RhoA, and the folate transporter RFC1 [2] (**Figure 8**), while also reducing DNA methylation [1], although the increased expression of CTR1, RhoA, and RFC1 did not correlate with changes in their promoter methylation. Overall, this suggests that decitabine could possibly reduce resistance to chemotherapy by increasing expression of transporters that are required for drug uptake, but that this increased expression of drug transporters is not directly related to reduction in promoter methylation.



To help guide our studies of factors associated with chemotherapy resistance in NSCLC, we have also published an extensive review on this topic [3]. Overall, factors associated with resistance may be associated with decreased drug delivery, uptake or activation, drug efflux pumps or detoxifying mechanisms, altered drug targets or DNA repair, or impaired apoptotic response or enhanced antiapoptotic response.

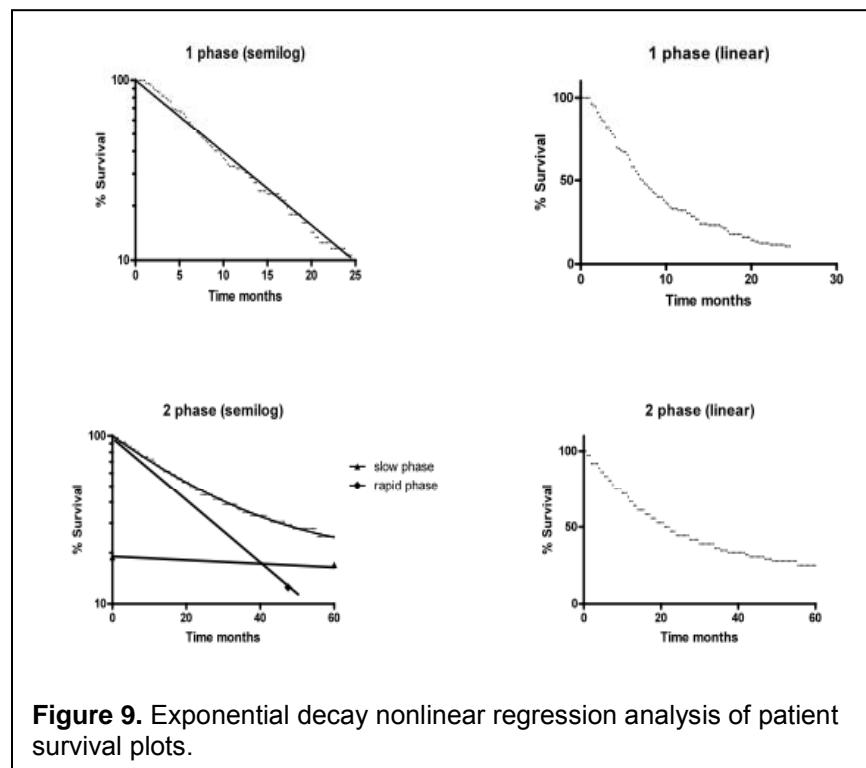


Figure 9. Exponential decay nonlinear regression analysis of patient survival plots.

related to discontinuation of a partially effective therapy after a standard period of treatment time, resulting in “synchronization” of patient deaths, and further hypothesized that assessing the characteristics of patients dying at times indicated along the leading edge of the convexity curve might identify a subpopulation that would benefit from continuation of therapy beyond usual stopping points.

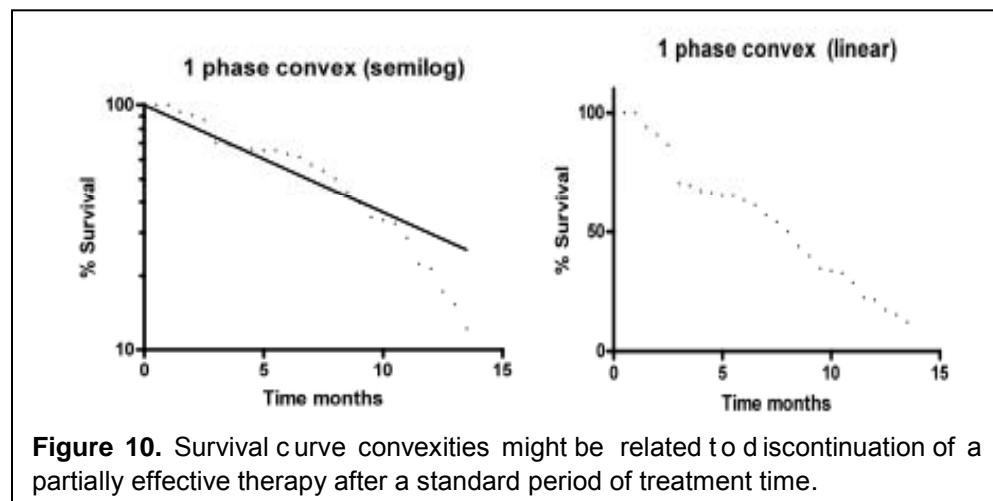


Figure 10. Survival curve convexities might be related to discontinuation of a partially effective therapy after a standard period of treatment time.

fraction (31%) consisting of patients with a half-life time to relapse of 24 months, and a rapidly-relapsing fraction (7%) of patients with a half-life time to relapse of 6 months. Using this type of approach, one can calculate, at each time point, what proportion of patients who are relapse-free but destined to relapse at some point in the future.

The overall survival curve for patients undergoing surgery was uniphasic, suggesting that co-morbidities and aging are also having a major impact on survival, and that assessment of time

We also proposed in PROSPECT to use exponential decay nonlinear regression analysis of patient survival plots to try to gain additional insights into factors affecting survival, and a manuscript is now in press on preliminary work using this approach [4]. Examples of how curve shapes differ in linear Kaplan-Meier plots vs log-linear exponential plots are illustrated in **Figure 9**. We also found that some types of therapies were associated with convexities on the log-linear plots. We hypothesized that these survival curve convexities (**Figure 10**) might be

Our initial assessments also suggested that following resection of stages I-III NSCLC, there is a true “cured” fraction (61%) of patients with a half-life time to relapse of greater than life expectancy, a slowly-relapsing

to relapse would give a better assessment of the impact of tumor markers on prognosis than would overall survival curves.

Key Research Accomplishments

- Collected tumor specimens from 457 lung cancer patients (including 117 who had neoadjuvant chemotherapy).
- Collected blood samples from 461 lung cancer patients (from whom tumor is also available in 345).
- Performed preliminary correlations of several biomarkers with survival and with stage and tumor type, as outlined in more detail in the Project 3 report.
- Performed preliminary assessments of the use of exponential decay nonlinear regression analysis of patient survival curves.
- Defined effect of the demethylating agent decitabine and of time from last therapy on expression, in human tumors, of transporters that may play a role in chemotherapy uptake.
- Identified numerous potential resistance mechanisms that will be investigated further.
- Correlated patient survival and time to recurrence with per cent tumor viability post - neoadjuvant chemotherapy.

Conclusions

During this project period, we have completed collection of the planned specimens ahead of schedule and have made good progress on preliminary assessment of these specimens. Molecular profiles will be generated from these specimens and correlated with patient outcomes and resistance to chemotherapy.

Project 3: Molecular Profiling of Non-Small Cell Lung Cancer Tissue Specimens and Serum and Plasma Samples: Correlation with Patient Response and Tumor Resistance to Chemotherapy.

(Leader: Dr. Ignacio Wistuba; Co-Leaders: Lin Ji and John Minna)

Hypothesis:

In Project 3, we hypothesize that systematic molecular profiling of surgically resected non-small cell lung cancer (NSCLC) tissue specimens using therapeutic target-focused (TTF) and mRNA approaches, along with serum phosphopeptide screening and plasma DNA analysis, will lead to the following results:

1. Validation in patients' tissue specimens of molecular signatures obtained from NSCLC cell lines that are associated with *in vitro* and *in vivo* (xenograft) resistance of NSCLC cell lines to chemotherapeutic and targeted agents.
2. Identification of molecular profiling signatures associated with NSCLC sensitivity or resistance to chemotherapeutic agents that can identify NSCLC patients most likely to respond to a given targeted therapeutic agent.
3. Development and validation of serum phosphopeptide profiles and plasma DNA markers associated with NSCLC patient response and tumor resistance to chemotherapeutic agents.

Objectives:

The greatest obstacle to creating effective treatments for lung cancer is the development of resistance to both chemotherapeutic and targeted agents. In this highly integrated and

translational program project, we tackle one of the most clinically significant problems in lung cancer: the prediction of patient response to therapy, especially in the context of tumor resistance to current standard chemotherapies. The main objectives of this project are as follows:

- a) To profile surgically resected tumor tissue specimens obtained from NSCLC patients to validate molecular signatures found in the TTF and mRNA profiles developed in Project 1. These profiles will be compared with molecular signatures obtained from NSCLC cell lines that are associated with *in vitro* and *in vivo* (xenograft) resistance to chemotherapeutic and targeted agents.
- b) By comparing NSCLC tumor specimens (collected in Project 2) from patients who have received preoperative chemotherapy and from those who have not, to validate TTF and mRNA signatures that are found in Project 1 to be associated with resistance to therapy and with the activation of resistance-associated molecular pathways or that are found in Project 1 to be potentially exploitable as new therapeutic targets.
- c) To identify serum and plasma biomarkers as surrogate markers to predict the response of NSCLC patients to neoadjuvant chemotherapy and to predict patient outcome.
- d) To provide tissue- and serum-based molecular profile signatures or markers to Project 2 that can predict the clinical outcome of NSCLC patients who had undergone surgical resection with curative intent, with or without neoadjuvant therapy.

This interdisciplinary research proposal for profiling cell lines, tumor tissue, and serum samples from NSCLC patients requires extensive histopathological, molecular, and immunohistochemical studies, which will be coordinated and/or performed by the Pathology Core (see Pathology Core's report).

Specific Aims:

Aim 1: To validate, in retrospectively collected NSCLC tumor tissue specimens, the TTF and mRNA profiles predictive of the *in vitro* and *in vivo* (xenograft) resistance of NSCLC cell lines to chemotherapeutic and targeted agents.

Summary of proposal: We will select 150 surgically-resected NSCLC tumor specimens from The University of Texas Lung SPORE (UT-SPORE) Tissue Bank for TTF and mRNA profiling. Using those 150 frozen archival NSCLC tumor tissues, we will perform reverse-phase protein array (RPPA), multiplex bead-based protein analysis (MBA) and Affymetrix U133 Plus 2.0 array to validate the molecular signatures developed in Project 1. Then, we will compare the profile signatures obtained from the NSCLC tumor specimens with the signatures obtained from NSCLC cell lines in Project 1 that predict the *in vitro* and *in vivo* resistance to chemotherapeutic and targeted agents. Finally, using formalin-fixed and paraffin-embedded tissue specimens, we will validate the expression of proteins abnormally represented in the molecular profiling analyses of NSCLC tumor specimens by using tissue microarrays (TMAs) and semiquantitative immunohistochemical (IHC) methods.

Summary of Research Findings

During the third year of research, in collaboration with the Pathology Core (I. Wistuba), we have performed a comprehensive molecular profiling of NSCLC and MPM tissue specimens with annotated clinical information, and achieved the following milestones (Figure 1):

- 1) Finalized the selection of 209 surgically resected, chemo-naïve NSCLC tumors with annotated clinicopathologic data, including adjuvant therapies and outcomes.
- 2) Completed the mRNA profiling (Illumina array platform) of 209 NSCLC tumors.

- 3) Expanded and completed the molecular profiling of NSCLC tumors to include miRNA (Agilent platform; n=209 cases).
- 4) Expanded and completed the DNA copy number analysis (array Comparative Genomic Hybridization, aCGH; n=162 cases).
- 5) Initiated the mutation analysis of genes frequently mutated in NSCLC (*KRAS* and *EGFR*).
- 6) Initiated the process for proteomic analysis using reverse phase protein analysis (RPPA) of 170 NSCLC tumors.
- 7) Designed a strategy for gene (mRNA), miRNA, and protein signature validation using high-throughput Quantitative Nuclease Protection Assay (qNPA™), Fluidigm™ microfluidic quantitative dynamic array, and IHC methodologies.
- 8) Completion of the manuscript describing the study of alternative approaches for mRNA profiling of NSCLC tissue specimens using FFPE samples.
- 9) In collaboration with Project 4 (A. Tsao), we have completed comprehensive molecular profiling of malignant pleural mesothelioma (MPM) tissue specimens and cell lines.

The detailed progress update is as follows:

1. Selection of 209 surgically resected NSCLC tissue specimens for molecular profiling as proposed in Aim 1. In collaboration with the Pathology and Biostatistics/Bioinformatics Cores, we finalized the selection of 209 surgically resected NSCLC tumors with annotated clinicopathologic data, including adjuvant therapy and outcome (**Table 1**). The final selection was performed using the following criteria: a) adenocarcinoma and squamous cell carcinoma histology; b) data available on adjuvant therapy; c) frozen tumor tissue with $\geq 70\%$ tumor content in the histology quality control; d) frozen tumor tissue with $\geq 30\%$ of malignant cell content; and, e) mRNA integrity number (RIN) ≥ 4 . The 209 cases selected were randomly selected by the Biostatistics/Bioinformatics Core (K. Coombes) for aliquoting of RNA and DNA and distribution by the Pathology Core to our Project 3 labs (**Figure 11**).

Table 1. Clinicopathologic characteristics of 209 chemo-naïve NSCLC profiled in Aim 1.			
Features	Adenocarcinoma	Squamous Cell Ca	Total
Number	152	57	209
Gender			
Male	74	20	94 (45%)
Female	78	37	115 (55%)
Path Stage			
I - II	124	47	171 (82%)
III - IV	28	10	38 (18%)
Smoking			
Yes	17	0	17 (18%)
No	135	57	192 (92%)
Adjuvant Chemotherapy			
Yes	69	29	94 (45%)
No	83	28	115 (55%)

2. mRNA profiling. Using one aliquot of RNA provided by the Pathology Core, Dr. Minna's lab has completed the mRNA profiling of the 209 NSCLC tumor specimens using the Illumina Array platform. The array data are currently under analysis to generate gene expression signatures that predict survival benefit for patients treated with platinum doublet-based adjuvant chemotherapy. Using *in vitro* drug response phenotype data on 100 NSCLC lines for platinum and taxane drugs alone and in combination derived from Project 1, coupled with these Illumina mRNA expression data, we will

also generate gene expression signature panels that predict cell line response to these drugs alone and in combination. These cell line predictive signatures will then be tested in the data obtained from the 209 NSCLC tumors. The analysis of 115 NSCLC specimens from patients

who did not receive adjuvant therapy will allow us to identify the predictive signatures that may also have prognostic value in that patient group.

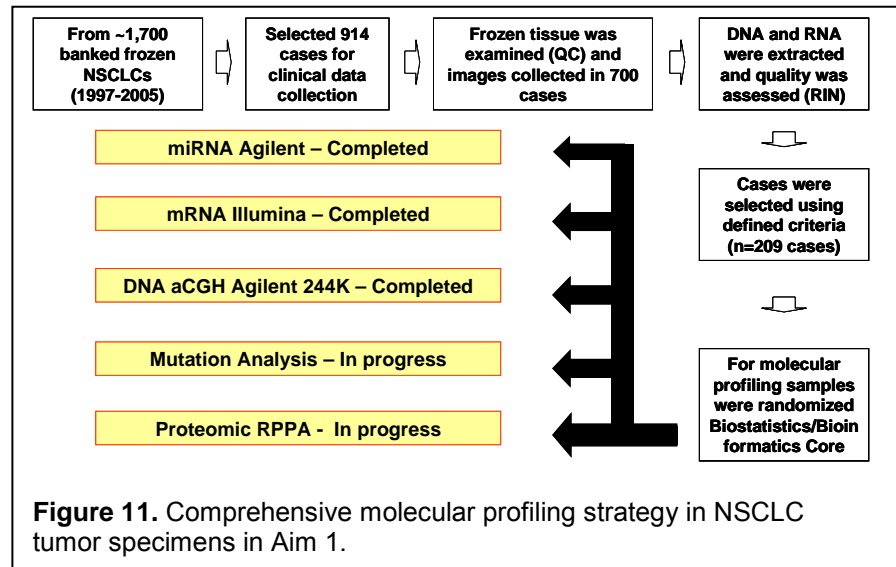
3. miRNA profiling. Using one aliquot of RNA provided by the Pathology Core, Dr. Wistuba's lab has completed the miRNA profiling using the Agilent v12 Human Array platform of the 209

NSCLC tumor specimens. The array data are currently under analysis by the Biostatistics/Bioinformatics Core to generate miRNA expression signatures that predict survival benefit for platinum doublet-based adjuvant chemotherapy. The analysis of NSCLC from patients who did not receive adjuvant therapy will allow us to determine if the predictive miRNA signatures may also have prognostic impact in those patients without adjuvant treatment. In addition, the integration of the mRNA with miRNA expression data will allow us to understand the biological interactions between mRNA and miRNA expression in NSCLC, and the identification of potentially novel pathways activated in NSCLC that might be regulated by miRNAs.

4. aCGH profiling. Using one aliquot of DNA provided by the Pathology Core, Dr. Wistuba's lab has completed the DNA copy number profiling using the Agilent 244K aCGH platform on 162 NSCLC tumor specimens with high content of malignant cells in the specimens. The aCGH data are currently under analysis by the Biostatistics/Bioinformatics Core to generate copy number signatures that predict survival of patients who did not receive and received platinum doublet-based adjuvant chemotherapy. The analysis of NSCLC from patients who did not receive adjuvant therapy will allow us to determine if the potentially predictive DNA copy number signatures also may have prognostic impact in the untreated patients. In addition, the integration of the DNA copy number data with the mRNA and miRNA data will allow us to explore potential novel interactions between gene (mRNA) and miRNA expression mechanisms.

5. DNA mutation analysis. Using one aliquot of DNA provided by the Pathology Core, Dr. Wistuba's lab is performing the analysis of *KRAS* (codons 12, 13 and 61) and *EGFR* (exons 19 and 21) status using PCR-base sequencing methodology in all 209 NSCLCs used for profiling. To date, from 227 NSCLC cases tested, *KRAS* mutations have been detected in 40 out of 136 (29%) adenocarcinomas and in none of 91 squamous cell carcinomas examined. This work will be completed during the next year of the grant. Additionally, one aliquot of DNA has been prepared and reserved for mutation analysis of a 20-gene panel using the MALDI-TOF Mass Spectrometry-based (SNP) analysis, Sequenom®.

6. Proteomic RPPA profiling. One frozen tissue specimen from each of the 170 NSCLC tumors, which are part of the 209 cases selected for molecular profiling, were distributed by the Pathology Core to Dr. Heymach's lab for protein extraction and RPPA analysis of a 176-protein panel. This work will be performed and completed during the next year of the grant.



7. Alternative approaches for mRNA profiling using FFPE NSCLC samples. As reported last year, we have explored alternative approaches for the molecular profiling of tissue specimens. In the study that was reported last year, we tested the feasibility of developing lung cancer prognosis gene signatures using genome-wide expression profiling of FFPE samples, which are widely available and provide a valuable rich source for studying the association of molecular changes in cancer and associated clinical outcomes. We randomly selected 100 NSCLC FFPE samples with annotated clinical information, microdissected the tumor area from the FFPE specimen, and then used Affymetrix U133 plus 2.0 arrays to attain gene expression data. After strict quality control and analysis procedures, a supervised principal component analysis was used to develop a robust prognosis signature for NSCLC. Three independent published microarray data sets were used to validate the prognosis model. This study demonstrated that the robust gene signature derived from genome-wide expression profiling of FFPE samples is strongly associated with lung cancer clinical outcomes and can be used to refine the prognosis for stage I lung cancer patients, and is independent of clinical variables. This signature was validated in several independent studies and was refined to 59-gene lung cancer prognosis signature. We concluded that genome-wide profiling of FFPE lung cancer samples can identify a set of genes whose expression level provides prognostic information across different platforms and studies, which will allow its application in clinical settings. A manuscript detailing this work was completed in the past year and recently submitted to the *Journal of Clinical Oncology*.

8. Molecular signatures validation strategies. Quantitative Nuclease Protection Assay (qNPA™) using FFPE tumor tissue specimens. We have designed a strategy for gene (mRNA), miRNA, and protein signature validation using the following methodologies: a) Quantitative Nuclease Protection Assay (qNPA™; for mRNA and miRNA signatures); b) Fluidigm™ microfluidic quantitative dynamic array; and, c) IHC. In collaboration with High-Throughput Genomic (HTG), Inc. (<http://www.htgenomics.com/technology/qnpa>), and the Pathology Core, we are developing multiple qNPA™ assays to test the expression of up to 200 genes using both frozen and FFPE tissues from surgically resected NSCLC tumors. Currently, FFPE tissue samples from all 209 NSCLC are being processed by the Pathology Core and RNA lysates being prepared for qNPA™ analysis. Preliminary analysis performed in NSCLC cell lines and FFPE tissue specimens have demonstrated the feasibility of this approach. Using the Fluidigm™ methodology, we have tested mRNA expression of 86 cancer stem cell markers and 10 endogenous control genes using 20 ng of total RNA per sample in RNA extracted from 80 NSCLC frozen tissues (**Figure 12**). Currently, we are preparing pre-amplified total RNA from matched FFPE samples for Fluidigm™ analysis using archival tissue specimens.

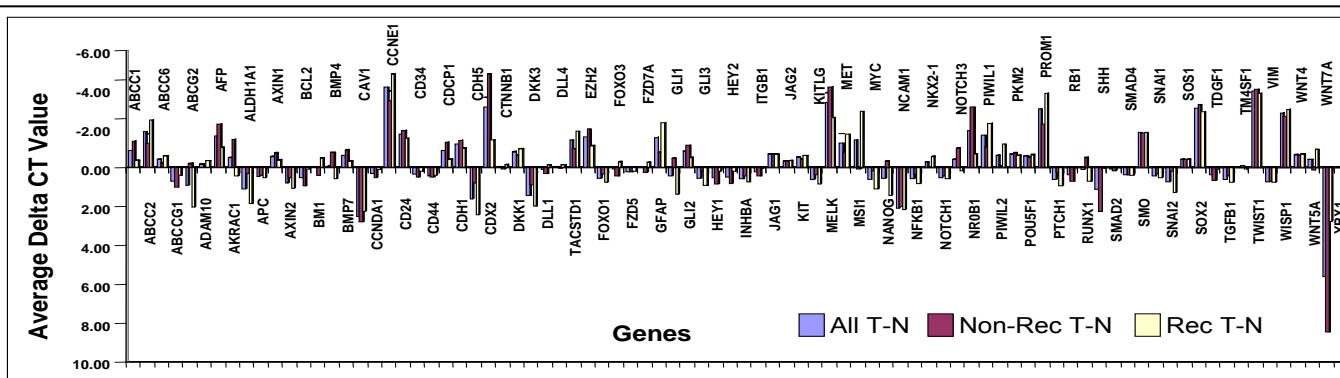


Figure 12. Differential mRNA gene expression levels using Fluidigm™ in 80 NSCLC tumor specimens comparing tumor (T) vs. normal (N) in all cases, and tumors with recurrence (n=40) and without recurrence (n=40) after 3 years follow-up. Negative Average Delta CT values indicate gene overexpression, and positive values loss or reduction of expression.

9. Comprehensive molecular profiling and bioinformatic analysis of MPM tissue and cell lines specimens. In collaboration with Project 4 (A. Tsao), we have performed a comprehensive profiling and bioinformatic analysis of 53 MPM tissue specimens with paired controls and 5 MPM cell lines. These studies have been led by Dr. M. Suraokar in Dr. Wistuba's lab.

a) *MPM tissue messenger RNA Profiling and Bioinformatic Analysis Study.* We had extracted total RNA from 91 tissue samples, representing 53 cases of MPM [including 38 cases with

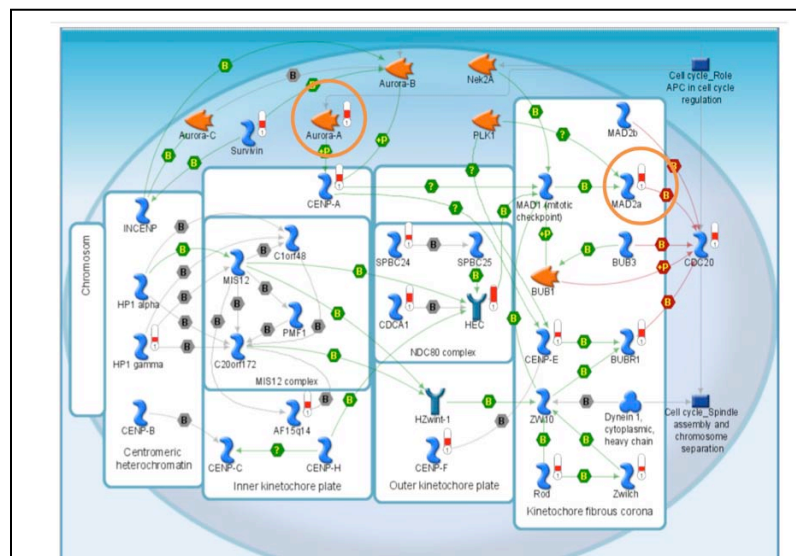


Figure 13. Pathway analysis using MetaCoreSuite (GeneGo Inc.) shows Mitotic Spindle Assembly Checkpoint (MSAC) pathway map with physical locations of 15 components upregulated (red bars) in pleural mesothelioma tumors including *Mad2a* message - circled (3 fold high at $p=4.12e-11$) and *AURKA* message - circled (3.4 fold high at $p=1.24e-12$).

vs. normal samples. About ~1950 highly significant probe sets representing ~670 genes, at a False Discovery Rate of $1e-09$, were obtained and subjected to pathway analysis using MetaCore software suite (GeneGo, Inc.). The most significantly altered pathway in MPM tumors was the Mitotic Spindle Assembly Checkpoint (MSAC) pathway due to up-regulation (greater than 2-fold) of at least 15 genes (**Figure 13**). These genes include mitotic arrest deficient-like 1 (*Mad2L1/Mad2a*), survivin (*BIRC5*), and therapeutic targets like

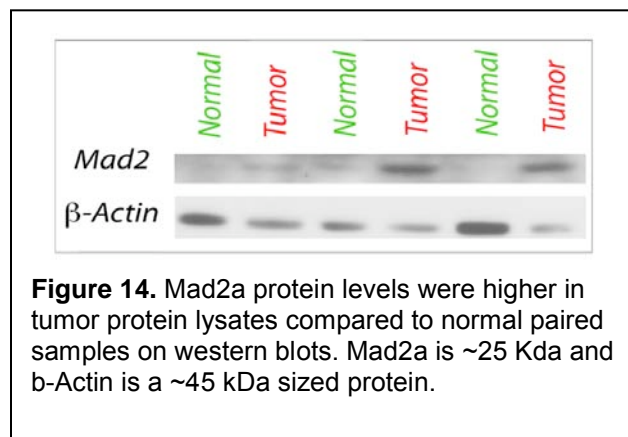


Figure 14. *Mad2a* protein levels were higher in tumor protein lysates compared to normal paired samples on western blots. *Mad2a* is ~25 Kda and β -Actin is a ~45 kDa sized protein.

paired-normal controls (adjacent and non-tumor tissue)] with epithelioid ($n=35$), biphasic ($n=13$) and sarcomatoid ($n=5$) MPM histology types. About 68% of the samples had greater than 70% tumor content and the same percentage yielded high quality RNA with RIN values ≥ 5 . Affymetrix U133 plus 2.0 microarray data was obtained for these samples at the UTMDACC Microarray Core Facility and analyzed by the

Biostatistics/Bioinformatics Core. Analysis of the microarray data using a 2-sample t-test was applied on a probe-by-probe basis followed by Beta-uniform Mixture for multiple comparisons. Finally, a paired t-test was applied to determine the differences between tumors

vs. normal samples. About ~1950 highly significant probe sets representing ~670 genes, at a False Discovery Rate of $1e-09$, were obtained and subjected to pathway analysis using MetaCore software suite (GeneGo, Inc.). The most significantly altered pathway in MPM tumors was the Mitotic Spindle Assembly Checkpoint (MSAC) pathway due to up-regulation (greater than 2-fold) of at least 15 genes (**Figure 13**). These genes include mitotic arrest deficient-like 1 (*Mad2L1/Mad2a*), survivin (*BIRC5*), and therapeutic targets like. Other published profiling studies on messenger RNA differentially expressed in mesothelioma tumors have showed different components of this pathway to be up regulated. However, our study shows for the first time a significantly high deregulation of this pathway.

We are currently cross-validating the expression of these Aurora kinase A (*AurKA*) genes using quantitative Polymerase Chain Reaction (PCR) platform. Messenger RNA profiling followed by bioinformatic analysis suggest that key MSAC genes are upregulated in 4 mesothelioma cell lines

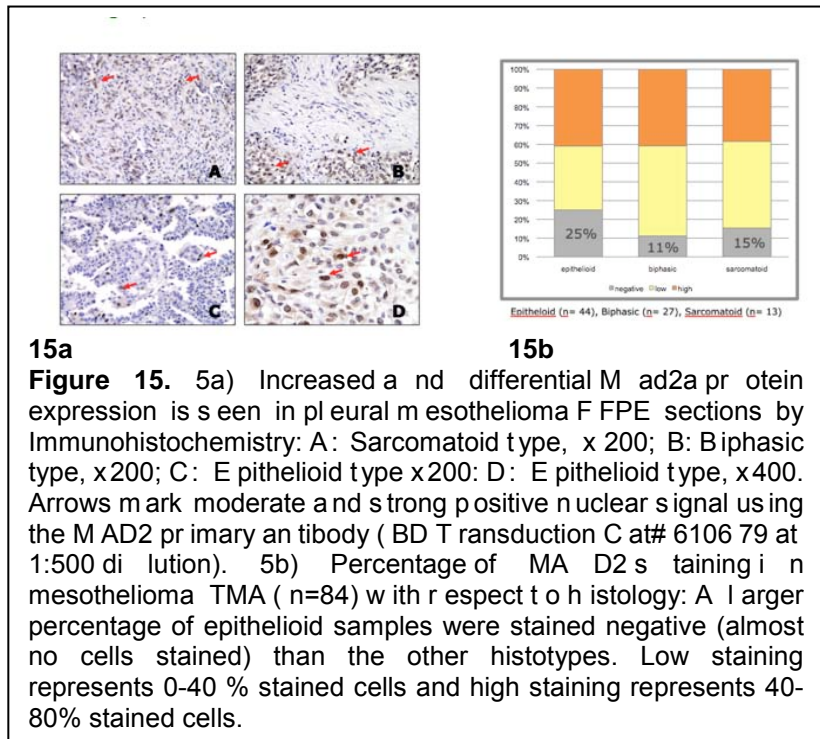
(H28, M STO-211H, H2052, and H 2452) compared to a control cell line (HCT-4012). These genes would serve as good controls for qPCR experiments and subsequent translational studies. Western Blot analysis of these tumor lysates, compared to normal controls, has confirmed the upregulation of the Mad2a protein (Figure 14).

Importantly, these protein lysates are obtained from the same tissue samples used for mRNA profiling and can be accurately correlated with the actual Mad2a message levels in the profiled tumors. Additionally, IHC of Mad2a protein in a separate TMA set

of 84 MPM tissues has also confirmed upregulation of this protein in tumors (Figure 15). We also see upregulation of other MSAC genes like AURKA and survivin protein in IHC sections of

this TMA. Currently, we are analyzing this observed upregulation using IHC, and will quantify expression using Western Blot analysis.

Interestingly, further bioinformatic analysis of the messenger RNA microarray data using probe sets, which were different between normal and tumor samples, defined 3 subtypes or groups (Figure 16). Comparing all probe sets between these 3 groups, using ANOVA at FDR level of $5e-05$, resulted in ~ 680 genes differentially expressed. These 3 groups partially overlapped with the histological subtypes – e.g., epithelioids were present in differing amounts between the 3 groups, but were mostly represented within group 1 (all epithelioids except for one biphasic sample). Group 2 had most of the biphasic samples while Group 3 had most of the sarcomatoid samples. MSAC pathway genes were differentially expressed between them. The sarcomatoid-rich Group 3



15a

15b

Figure 15. 5a) Increased and differential Mad2a protein expression is seen in pleural mesothelioma FPE sections by Immunohistochemistry: A: Sarcomatoid type, x200; B: Biphasic type, x200; C: Epithelioid type x200; D: Epithelioid type, x400. Arrows mark moderate and strong positive nuclear signal using the MAD2 primary antibody (BD Transduction Cat# 610679 at 1:500 dilution). 5b) Percentage of MAD2 staining in mesothelioma TMA (n=84) with respect to histology: A larger percentage of epithelioid samples were stained negative (almost no cells stained) than the other histotypes. Low staining represents 0-40% stained cells and high staining represents 40-80% stained cells.

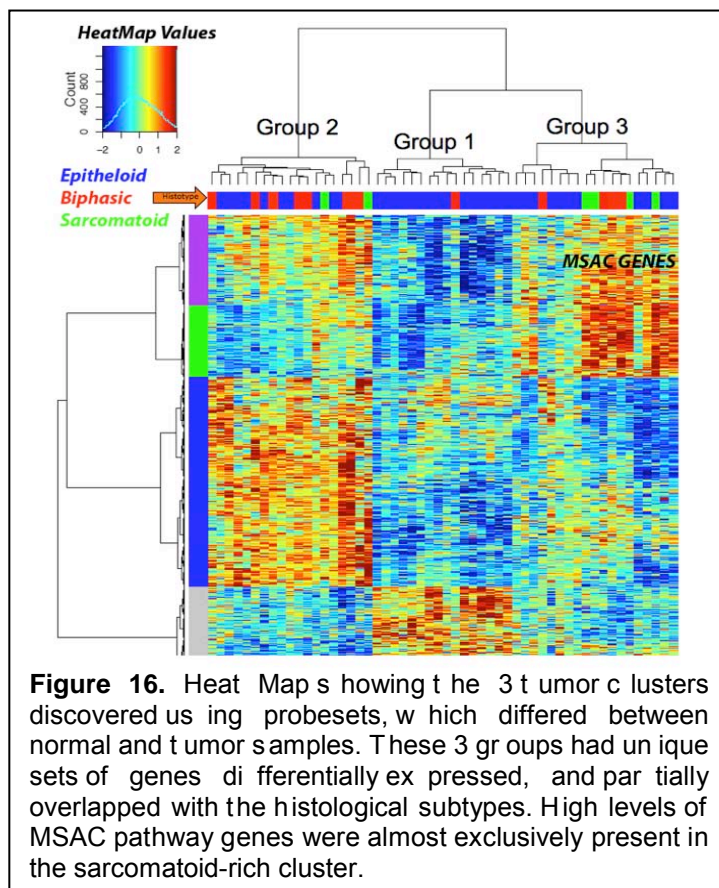
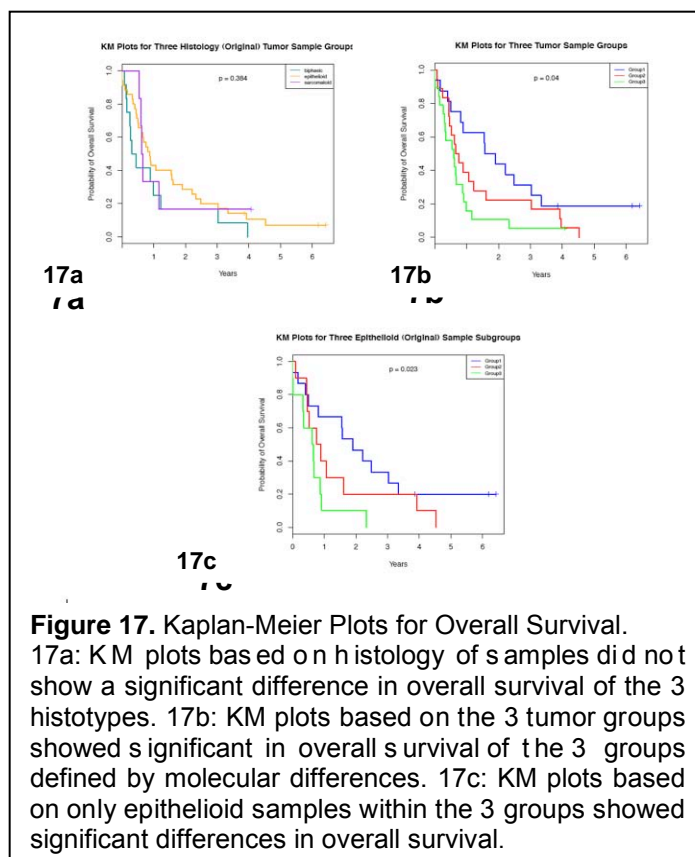


Figure 16. Heat Maps showing the 3 tumor clusters discovered using probesets, which differed between normal and tumor samples. These 3 groups had unique sets of genes differentially expressed, and partially overlapped with the histological subtypes. High levels of MSAC pathway genes were almost exclusively present in the sarcomatoid-rich cluster.

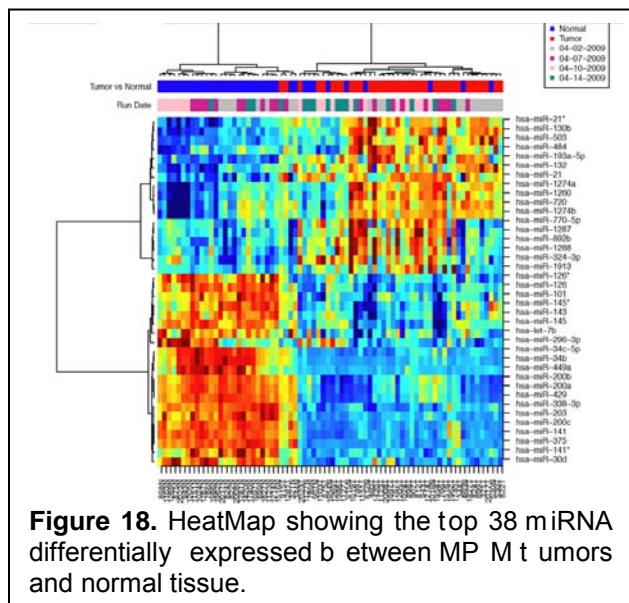
showed the highest levels of these genes. There are also a group of genes expressed at higher levels in epithelioid-rich group 1 and biphasic-rich Group 2. We are currently doing pathway analysis on these genes using Ingenuity Pathway Analysis and (Ingenuity, Inc.) and MetaCore suite. Our preliminary analysis has shown the presence of high levels of PTEN tumor suppressor pathway within Group 1 compared to the other 2 groups. PTEN has been reported to be an indicator of good prognosis in mesothelioma, which is consistent with our data.

Moreover, Kaplan-Meier (KM) plots for overall survival based on histology or the groups defined by molecular signatures have revealed a surprising result. It is a well-known observation that mesothelioma patients show differential overall survival based on the histology of the tumor. Those patients with sarcomatoid type MPM show the worst overall survival, while epithelioid-type MPM patients show relatively better chances of survival. The biphasic tumors affect the survival of patients to a degree midway between these 2 tumor types. However, we observe that the KM plots based on the 3 groups ($p=0.04$) defined by the messenger RNA molecular profile is a better indicator of overall survival of the patients compared to using the histological criterion ($p=0.384$) (**Figure 17**). This is a significant finding in the field of mesothelioma and provides an additional criterion to estimate overall survival of these patients. Moreover, this difference is not driven by the histotype in these 3 groups, because even the epithelioid tumors across these 3 groups show significant differences in overall survival ($p=0.023$). It has been reported in at least 2 different studies that epithelioids can be separated into 2 different groups – short-term and long-term survivors. Our observation confirms the heterogeneous nature of epithelioid tumors with respect to overall survival and, additionally, discovers a molecular basis for this difference. We are continuing these studies by pursuing a multivariate analysis of these cases based on known prognostic factors like histology, gender, age, stage, and treatment.



b) MPM MicroRNA profiling. The same set of 91 samples, which were profiled for messenger RNA, was also profiled in Dr. Wistuba's lab for microRNA content using the Agilent human miRNA microarray (Agilent Technologies, Inc., Santa Clara, California, USA). The microarray data is currently being analyzed by the Biostatistics/Bioinformatics Core. Their preliminary analysis suggests that a number of miRNAs are differentially expressed between the tumor and normal samples. These miRNAs can be clustered into 2 groups – one composed of all normal samples, and the other containing all tumor samples with a few normal samples. We are currently analyzing the reason for this discrepancy since our messenger RNA profiling on these same samples showed clear differences between the normal and tumor samples. **Figure 18** shows the top 38 miRNA's differentially regulated between these 2 groups, which were obtained by performing a

paired T-test on a miRNA-by-miRNA basis and were selected at a highly significant FDR value of $1e-06$. Many of these miRNAs agree with published reports from other groups especially the down-regulation of miR-200 cluster seen in tumors. Our next step is to validate the levels of these miRNAs in tumor samples using qPCR platform.



protein lysates have been run on Western Blots for determining the levels of MSAC proteins (see above). The protein lysates will be printed with the help of Dr. John Heymach's lab to obtain RPPA for proteomic analysis.

e) MPM Cell line Profiling Update. We have generated a unique and extensive collection of 21 mesothelioma cell lines, acquired from various sources, including 17 pleural mesothelial and mesothelioma cell lines of different histotypes as well 3 peritoneal mesothelioma and 1 peritoneal primary cell lines. These cell lines have been characterized as authentic mesothelioma cell lines using IHC with 7 different markers - Cytokeratin 5/6, calretinin, mesothelin, CEA (carcinoembryonic antigen), B72.3, CD15 (LeuM1), and TTF-1 (thyroid transcription factor-1). The first 3 markers positively identify mesothelioma cell lines as they are frequently upregulated in mesothelioma tumors, whereas the others are rarely seen and serve as negative controls. Also, in accordance with NIH regulations, all of these cell lines have been DNA fingerprinted and authenticated using STR loci PCR analysis in the M. D. Anderson Characterized Cell Line Core Facility.

We have comprehensively profiled 5 of these cell lines using multiple platforms – miRNA analysis using the Agilent Human microRNA microarray version 3.0 (Agilent Technologies, Inc., Santa Clara, California, USA) in Dr. Wistuba's lab, messenger RNA analysis using Affymetrix U133 plus 2.0, and SNP/copy number analysis on the Affymetrix SNP 6.0 platform in the MDACC Microarray Core facility. Bioinformatic analysis using GeneSpring GX11 software (Agilent Technologies, Inc., Santa Clara, California, USA) on 4 mesothelioma cell lines (H28, MSTO-211H, H2052, and H2452) compared to control cell line (HCT-4012) has shown that the MSAC genes including Mad2a and AURKA are upregulated in the cancer cell lines. These cell lines would serve as good *in vitro* models to study the effects of therapeutic drugs targeting the MSAC pathway in pleural mesothelioma. Additionally, these and other cell lines have also been profiled for their proteomic content on RPPAs in Dr. Heymach's laboratory (Project 1). Together, these multi-omic studies will help us to discover novel loci of critical importance in pathogenesis

c) DNA profiling. DNA was extracted, and after estimation and quality control, was analyzed for SNP and copy number variations using the Human 1M-duo platform (Illumina, Inc., San Diego, CA, USA), which contains more than 1 million SNPs along with copy number variation content. This data will be analyzed by the Biostatistics/Bioinformatics Core, and represent ~75 % of the tumor and normal samples analyzed previously for the global mRNA and miRNA content.

d) Protein Profiling. We have extracted proteins from 90% of these tumor and normal samples for Reverse Phase Protein Array (RPPA) analysis using the protocol obtained from Dr. Heymach's lab (Project 1). These

of mesothelioma. For instance, the DNA copy number analysis by Nexus 4.0 software (BioDiscovery, Inc.) of the MPM cell lines – MSTO-211H- showed amplification of the c-myc oncogenic locus, whereas the normal control cell line HCT-4012 lacked this genetic abnormality. The mRNA array data also showed increased number of transcripts from this locus and, more importantly, the RPPA array data showed increased c-myc protein expression in this mesothelioma cell line compared to normal (**Figure 19**).

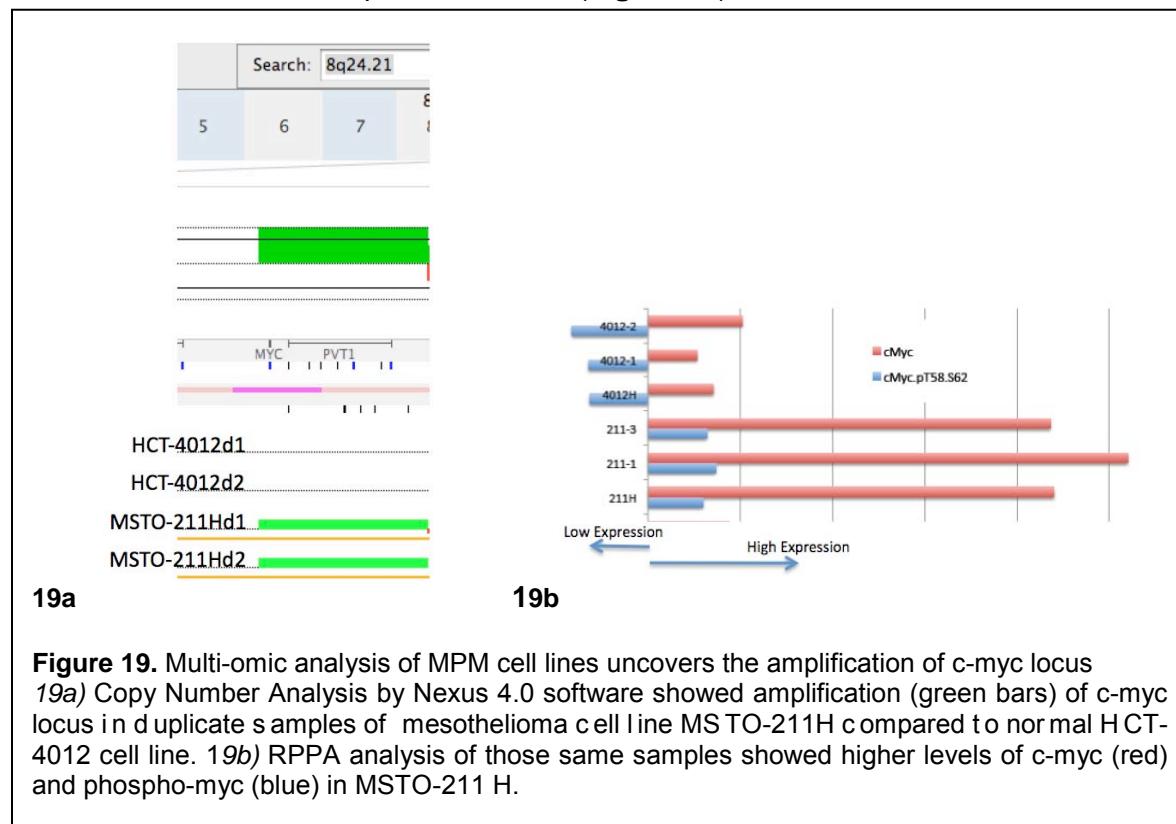


Figure 19. Multi-omic analysis of MPM cell lines uncovers the amplification of c-myc locus
19a) Copy Number Analysis by Nexus 4.0 software showed amplification (green bars) of c-myc locus in duplicate samples of mesothelioma cell line MSTO-211H compared to normal HCT-4012 cell line. 19b) RPPA analysis of those same samples showed higher levels of c-myc (red) and phospho-myc (blue) in MSTO-211H.

Aim 2: To develop TTF and mRNA signatures of NSCLC resistance to chemotherapy, and identify chemoresistance-associated targets/pathways as new therapeutic targets.

Summary of proposal: Whereas Aim 1 focuses on the identification in archived tumor specimens of TTF and mRNA molecular profiles detected in NSCLC cell lines, the main focus of Aim 2 is to determine whether the molecular signatures in the tumor specimens correlate with patient response to neoadjuvant chemotherapy. From the clinical trial in Project 2, we will use specimens from 100 NSCLC patients who received neoadjuvant therapy and had surgical resection with curative intent (cases) and from 200 NSCLC patients who had surgical resection but did not receive neoadjuvant therapy (controls) to perform RPPA, MBA, and Affymetrix U133 Plus 2.0 array analyses. Then, we will compare the TTF and mRNA profile signatures obtained from these NSCLC tumor specimens with signatures obtained in Project 1 to predict the *in vitro* and *in vivo* resistance of NSCLC cell lines to therapy. Those data will be provided to Project 2 for correlation with clinical characteristics, including prognosis and metastasis. Finally, using formalin-fixed and paraffin-embedded tissue specimens, we will validate the expression of proteins abnormally represented in the molecular profiling analyses in NSCLC tumor specimens from all patients enrolled in Project 2 by using TMAs and semiquantitative IHC methods.

Summary of Research Findings

During the third year, in collaboration with the Pathology Core (I. Wistuba), we have mainly

focused on the molecular profiling of surgically resected NSCLC obtained from patients who have received neoadjuvant chemotherapy. Our detailed progress update is the following:

- 1) Completion of comprehensive molecular profiling of 40 neoadjuvant-treated NSCLC using mRNA (Illumina), miRNA (Agilent), and DNA aCGH array profiling platforms.
- 2) Selection and processing for nucleic acids and protein extractions of additional 70 prospectively collected neoadjuvant-treated NSCLC cases.
- 3) Analysis of tissue and RNA quality of additional neoadjuvant-treated NSCLC tissue sets for profiling studies.
- 4) Completion of the data analysis and manuscripts submission of 3 additional projects reported last year, including: a) Expression of Keap1 and Nrf2 in NSCLC; b) Expression of cell membrane receptors in NSCLC; and, c) Expression of cancer stem cell markers in NSCLC.

1. Molecular profiling of neoadjuvant-treated NSCLC tumor tissues specimens. During the third year of the grant, we completed the mRNA (Illumina array; J. Minna), miRNA (Agilent array; I. Wistuba) and DNA aCGH (Agilent 244K; I. Wistuba) profiling of 40 neoadjuvant-treated tumors with annotated clinicopathologic data, including pathological response to therapy and patients' outcome. A frozen tissue specimen from each case has been received from the Pathology Core by Dr. J. Heymach's lab for protein extraction and RPPA analysis. The multiple platforms data are currently under analysis by the Biostatistics/Bioinformatics Core to generate signatures of chemotherapy resistant tumors and that predict survival of patients who received this type of preoperative therapy. The molecular signatures derived from these neoadjuvant-treated cases will be compared with those derived from the 209 chemo-naïve NSCLC tumors examined in Aim 1. In addition, the integration of the mRNA, miRNA, and DNA copy number array data obtained in the same set of tumors will allow us to explore potential novel regulations of gene expression (mRNA) and miRNA expression mechanisms in lung cancers resistant to chemotherapy. Similarly to chemo-naïve NSCLC tumors examined in Aim 1, the neoadjuvant-treated tumors are being tested for mutations in *KRAS* (completed) and *EGFR* (in progress).

2. Selection and processing for nucleic acids and protein extractions of additional 70 prospectively collected neoadjuvant-treated NSCLC cases. To achieve the goal of profiling 100 neoadjuvant-treated NSCLC tumors, we have selected 70 additional cases that have been recently processed for comprehensive molecular profiling by the Pathology Core. These studies will be completed during the fourth year of the grant.

3. Analysis of tissue and RNA quality of additional neoadjuvant-treated NSCLC tissue sets for profiling studies. As reported in previous years, Dr. Li Mao signed a collaborative agreement with the Intergroupe Francophone de Cancérologie Thoracique (IFCT) to obtain up to 250 frozen lung tumor tissues from patients enrolled in IFCT-0002 clinical trial (a open-labelled, multicenter, randomized phase III study), which was designed to define the best timing of neoadjuvant chemotherapies. The samples will be used to identify a gene expression signature of resistance to platinum-based chemotherapies. In collaboration with the Pathology Core, frozen tissue samples from 170 of these cases have been processed and assessed for histology quality control and RNA integrity. The preliminary data indicate that 95 neoadjuvant-treated NSCLC tumors showed $\geq 20\%$ of malignant cell content in the frozen tissue samples.

4. Completion of the data analysis and manuscripts submission of additional projects reported last year. During the previous years of the grant, we have developed additional projects to investigate in NSCLC novel biomarkers related or potentially related to resistance to chemotherapy. Four studies have been successfully completed, two manuscripts reporting the

main findings have been published, and one manuscript was recently submitted (see reported outcomes). A brief description of the major findings of each study is included below.

a) Expression of cancer stem cell markers in NSCLC. As previously reported, we studied the protein expression of a panel of 7 stem cell markers (EZH2, SOX2, CD24, CD44, C-kit, BMI-1, and Oct3/4) from a large series of chemo-naïve NSCLC (see abstract by Yuan et al presented at the WLCC 2010). We have hypothesized that stem cell markers could be involved NSCLC resistance to chemotherapy. One of these markers, SOX2 gene, was shown last year by the Dana-Farber Cancer Center group to be frequently amplified in squamous cell carcinoma of the lung, and considered a novel gene important in the pathogenesis of lung cancer. Our recently reported findings highlighted the cell-lineage gene expression pattern for the stem cell transcriptional factor SOX2 in the pathogenesis of lung squamous cell carcinoma, and suggested a differential activation of stem cell-related pathways between squamous cell carcinomas and adenocarcinomas of the lung. (Yuan et al, published in *PLoS One*, 2010). Data analysis of the remaining stem cell markers is in progress.

b) Expression of Keap1 and Nrf2 in NSCLC. Nuclear factor erythroid-2-related factor 2 (Nrf2) is a transcription factor associated with chemotherapy resistance and tumor growth, which is repressed by Kelch-like ECH-associated protein 1 (Keap1). We tested the hypothesis that the abnormal expression of these two proteins correlated with NSCLC patients' outcome and response to adjuvant chemotherapy. We demonstrated that increased Nrf2 expression and decreased Keap1 expression are common abnormalities in NSCLC and are associated with clinical outcome. In our study, abnormal expression of Nrf2 and Keap1 proteins was more common than that of the corresponding gene mutations, suggesting that other mechanisms are involved in the activation of NFE2L2 (Nrf2 gene) and inactivation of KEAP1. Nrf2 expression may play a role in response to adjuvant platinum-based chemotherapy in patients with squamous cell carcinoma. Identifying patients with abnormal Nrf2 expression may be important for selection for chemotherapy in NSCLC (Solis et al, published in *Clinical Cancer Research*, 2010).

c) Expression of cell membrane receptors in NSCLC. Membrane transporters FR α and RFC1 are potential biomarkers of tumor response to antifolate chemotherapy. Information on the protein expression of these receptors in NSCLC is limited. Here, we report for the first time that NSCLC frequently overexpressed FR α and RFC1 proteins by studying a large series of cases with annotated clinico-pathologic information. Importantly, we report that tumor cells from lung adenocarcinoma histology expressed significantly higher levels of cytoplasmic and membrane FR α than squamous cell carcinoma, and tumors from never-smokers were significantly more likely to express cytoplasmic FR α than those from smokers. In lung adenocarcinomas, the presence of *EGFR* mutations correlated with higher expression of membrane FR α and *FOLR1* gene expression. We postulate that this information may be useful in selecting which patients with NSCLC may benefit from and should receive treatment with antifolate agents, including pemetrexed.

d) VEGFR2 gene copy gain is predictive of shorter overall survival in NSCLC patients treated with platinum adjuvant chemotherapy. We studied the role of vascular endothelial growth factor-2 (VEGFR2) gene (*KDR*) abnormalities in malignant cells of surgically resected NSCLC tissues and correlated with patients' outcome after treatment with platinum adjuvant chemotherapy. We studied tissues obtained from 248 surgically resected NSCLCs. *KDR* copy number gain (CNG) was examined by quantitative PCR and fluorescence *in situ* hybridization. VEGFR2 protein expression and microvascular density were studied by immunohistochemistry. In NSCLC cell

lines, *KDR* CNG (n=75) and VEGFR2 levels (n=63) were quantified and correlated with *in vitro* sensitivity to platinum drugs. *KDR* mutation (exons 7, 11 and 21) and single nucleotide polymorphisms (SNPs; 889G/A, 1416A/T and -37A/G) were genotyped by PCR-based sequencing. Malignant cells demonstrated *KDR* CNG in 32% of NSCLC tumors. *KDR* CNG in malignant cells was associated with poor overall survival (OS) (HR=4.0; P=0.001) and worse recurrence-free survival (HR=1.83; P=0.044) in multivariate analysis. *KDR* CNG predicted worse OS (HR=5.16; P=0.003) in patients who received platinum adjuvant therapy but not in untreated patients (P=0.349). In cell lines, *KDR* CNG and high VEGFR2 expression correlated significantly with resistance to platinum. *KDR* mutations were not detected in NSCLC tumor tissues. The *KDR* variant genotypes SNPs 1416 AT/TT and -37 AG/GG were associated with a favorable OS in lung adenocarcinoma. The association between *KDR* CNG and worse outcome in platinum adjuvant therapy-treated NSCLC patients suggests that *KDR* might be a potential biomarker for predicting the efficacy of adjuvant chemotherapy in this disease.

Aim 3: To identify surrogate serum phosphopeptide profiles and plasma DNA markers associated with NSCLC tumor resistance and patient response to neoadjuvant chemotherapy.

We will identify serum samples from the UT-SPORE Tissue Bank that match the NSCLC tumor resection specimens examined in Aim 1. We will use these serum samples for phosphopeptide profiling and peptide mapping by ProteinChip array-based surface-enhanced laser-desorption/ionization (SELDI) mass spectrometry (MS) and laser desorption/ionization (LDI) mass spectrometry (MS)/MS to compare serum phosphopeptides with TTF and mRNA profiles. The phosphopeptide MS profiles from retrospective specimens will later be used as references and controls for the prospective serum proteomic analysis. As in Aim 2, we will use serum samples collected prospectively in Project 2 from 100 NSCLC cases undergoing neoadjuvant chemotherapy and 200 NSCLC controls undergoing surgery without neoadjuvant chemotherapy, and, when relevant, at the time of relapse. Using these serum specimens, we will perform phosphopeptide profiling on ProteinChip arrays by SELDI-MS to measure the temporal changes in serum phosphopeptides before and after the therapeutic intervention. We will use LDI-QSTAR-MS/MS and liquid chromatography (LC)-MS/MS to identify specific serum phosphopeptides that are determined by SELDI-MS to be relevant to targeted therapeutic response and acquired resistance in lung cancer patients. In addition, we will compare serum phosphopeptide profiles with TTF (RPPA and MBA) profiles, mRNA profiles, and TMAs and IHC analysis developed in Project 1 and in Aims 1 and 2 of this project. This comparison will identify TTF serologic molecular signatures and elucidate the biologic pathways potentially associated with patient response and tumor resistance to targeted therapeutic agents. Finally, in collaboration with Project 2 we will perform correlation analysis of these NSCLC serum phosphopeptide profile signatures with patients' clinical characteristics to predict lung cancer, cancer progression, cancer stages, and overall survival rate; to characterize serum phosphopeptide proteomic patterns and signatures in correlation to tumor recurrence, clinical response to adjuvant chemotherapeutic and targeted agents, and development of resistance; and to identify serum phosphopeptide markers as surrogate predictors of patient outcome.

Moreover, in Aim 3 we will quantify total circulating plasma DNA and methylation-specific DNA in all 300 patients with NSCLC enrolled in the Project 2 clinical trial. The circulating DNA levels will be correlated with patients' clinicopathologic characteristics. Any changes in these levels during chemotherapy and after surgery will be correlated with patient response to neoadjuvant therapy and patient outcome after surgery. The correlation between circulating methylated DNA levels and tumor DNA methylation will also be examined in a selected panel of patients.

Summary of Research Findings

Reversible protein phosphorylation is a key regulating switch that controls a wide range of biological functions and plays a critical role in signaling pathways involved in oncogenesis. We have developed an innovative functional proteomics platform using ProteinChip array-based SELDI-MS for high throughput profiling and identification of phospho-peptides in human serum to identify specific phosphopeptides/phosphoproteins associated with human lung cancer. We performed phosphopeptide profiling on serum samples from human normal and lung cancer patients with varying stages and smoking histories. We used phospho-tyrosine antibody-conjugated super-paramagnetic beads to capture phosphopeptides generated in trypsin-digested serum samples. The affinity-enriched pYPs were then randomly loaded onto SEND ProteinChips with duplicates and analyzed by SELDI-TOF Mass Spectrometry. We used wavelets and the mean spectrum for peak detection and detected more than 600 pYP peaks spanning a M/Z range from 50 to 5500 Dalton. For each peak, we recorded the p-value from an F-test and modeled the set of p-values using a beta-uniform mixture model to estimate the false discovery rate (FDR). We identified 39 pYP peaks with fold changes in intensity detected on SELDI-MS profiles to be significantly (at FDR = 10%) differentially expressed in normal and lung cancer serum samples. The phosphopeptides detected on SELDI-MS spectra were further identified using a protein chip array-interfaced qSTAR-MS/MS. One of the phospho-tyrosine containing peptides was identified as an Alpha-1-acid glycoprotein 1 precursor (A1AG1) or ORM-1 serum protein. The ORM-1 pYP showed a M/Z peak at 1752.3 Da and was significantly upregulated in lung cancer serum samples, with more than 10-fold increase ($P = 0.0024$) in mass peak intensity. A computer-aided structural and function analysis predicted the potential association of ORM-1 to the nicotinic acetylcholine receptor (nAChRs) and to be a substrate of EGFR tyrosine kinase.

We further validated phospho-ORM-1 protein expression in another set of 80 lung cancer and control serum samples by ELISA and confirmed the significantly upregulated expression of serum phospho-ORM-1 in late stage lung cancer patients with ever-smoking history. We also identified protein interactions between the ORM-1 and subunits of nAChRs in lung cancer cell lines by immunoprecipitation and immune-blotting analysis. Our results suggest the role of the Phospho-ORM-1 peptide as a novel NACHR-associated protein in lung cancer pathogenesis and smoking-associated carcinogenesis and as a potential serum marker for lung cancer detection. We are now validating our findings in another set of human lung cancer serum samples (80 samples consist of 20 for never-smoker controls, and 20 ever-smoker, 20 Stage III/IV, and 20 stage I/II lung cancer cases) and will perform phospho-peptide profiling in 150 matching serum samples that were used for other biomarker analyses in this project. These findings were presented at the 101th AACR Annual meeting in Washington D.C., 2010 (Abstract #1624.)

Key Research Accomplishments

- Completion of comprehensive molecular profiling (mRNA, miRNA, and DNA-aCGH) of a large series (n=209) of chemo-naïve NSCLCs with annotated clinicopathologic information, including outcome after adjuvant chemotherapy.
- Completion of molecular profiling (mRNA, miRNA, and DNA-aCGH) of 53 MPM tissue specimens and cell lines, and identification of molecular signatures associated with outcome of patients with surgically resected MPM and novel MPM potential molecular targets.
- Completion of comprehensive molecular profiling (mRNA, miRNA, and DNA-aCGH) of 40 chemotherapy neoadjuvant-treated NSCLCs with annotated clinicopathologic information, including patients' outcome.

- Characterization of NSCLC tissue specimens for novel biomarkers associated with resistance to chemotherapy in lung cancer, including Nrf2/Keap1 expression, *KDR* copy number, and membrane transporters expression.
- Identification and characterization of phospho-ORM-1 as a novel nicotinic acetylcholine receptor (NACHR)-associated protein and a potential serum marker for lung cancer detection.

Conclusions

During the third year, we reached the goal of molecular profiling chemo-naïve surgically resected NSCLCs from patients who received and did not receive adjuvant chemotherapy. Also, we have completed similar profiling in MPM tissue and cell line specimens, devised gene expression signatures associated with MPM patients' outcome, and identified novel potential molecular targets. In addition, we have characterized two novel markers (Nrf2 and *KDR*) associated with response to platinum-based adjuvant chemotherapy in NSCLC.

Project 4: Target Modulation Following Induction Treatment With Dasatinib in Patients With Malignant Pleural Mesothelioma (MPM) and Identification of New Therapeutic Targets/Strategies for MPM

(Leaders: Drs. Anne Tsao, Reza Mehran)

Hypothesis:

We hypothesize that dasatinib, a broad spectrum ATP-competitive inhibitor for oncogenic tyrosine kinases (BCR-ABL, SRC, c-Kit, PDGFR, and ephrin receptor kinases), may be a new therapeutic agent in malignant pleural mesothelioma (MPM). We also believe that conducting therapeutic target-focused (TTF) molecular and gene profiling (Affymetrix arrays) will lead to development of other novel therapies for MPM.

Specific aims:

Aim 1: Conduct a phase I clinical trial with the primary endpoint of biomarker modulation using dasatinib as induction therapy in patients with resectable MPM.

Trial design: Untreated MPM patients undergo extended surgical staging (ESS) with multiple biopsies to account for tumor heterogeneity. If deemed a surgical candidate for either P/D or EPP, patients receive 4 weeks of oral dasatinib (70 mg BID) followed by P/D or EPP. If a radiographic response is seen, an additional 2 years of dasatinib maintenance after adjuvant radiotherapy and chemotherapy is given. Serum/blood/platelets/pleural effusion specimens are collected for exploratory analysis of peripheral surrogate biomarkers. Primary endpoint is biomarker modulation of Src Tyr419 immunohistochemistry (IHC) in tumor tissue. Secondary endpoints: response, survival, safety/toxicity, and biomarker modulation.

- 1a. Determine the effects of dasatinib induction therapy on selected tumor biomarkers (activated Src, PDGFR, VEGFR) pre- and post-induction therapy.
- 1b. Determine the modulatory effects of dasatinib on selected biomarkers of survival and apoptosis (PI3K/AKT, bcl-xL, caspases), proliferation (IGFR, Ki-67), angiogenesis (IL-8, bFGF, TNF- α), epithelial-mesenchymal transition (TNF- β , E-cadherin, c-Kit/Slug) and invasion/migration (Ephrin, MMP) in tumor specimens pre- and post-induction therapy.
- 1c. Determine the effects of induction dasatinib therapy on tumor mean vessel density, cell apoptosis, and the proliferation index.

- 1d. Determine the modulatory effects of dasatinib on serum, platelet, and pleural effusion markers of survival (PI3K/AKT, bcl-xL, caspases), proliferation (IGFR, Src), angiogenesis (soluble VEGFR, VEGF, PDGF, IL-8, bFGF, TNF- α), and invasion/migration (Ephrin, MMP).
- 1e. Determine the drug concentration of dasatinib in tumor and serum.
- 1f. Assess the effects of dasatinib and cytoreductive surgery on the serum mesothelin-related peptide (SMRP) level.
- 1g. Assess the safety and toxicity profile of induction dasatinib in patients with resectable MPM.

Aim 2: Conduct radiographic correlates of tumor response and clinical outcome with positron-emission technology-computer tomography (PET-CT).

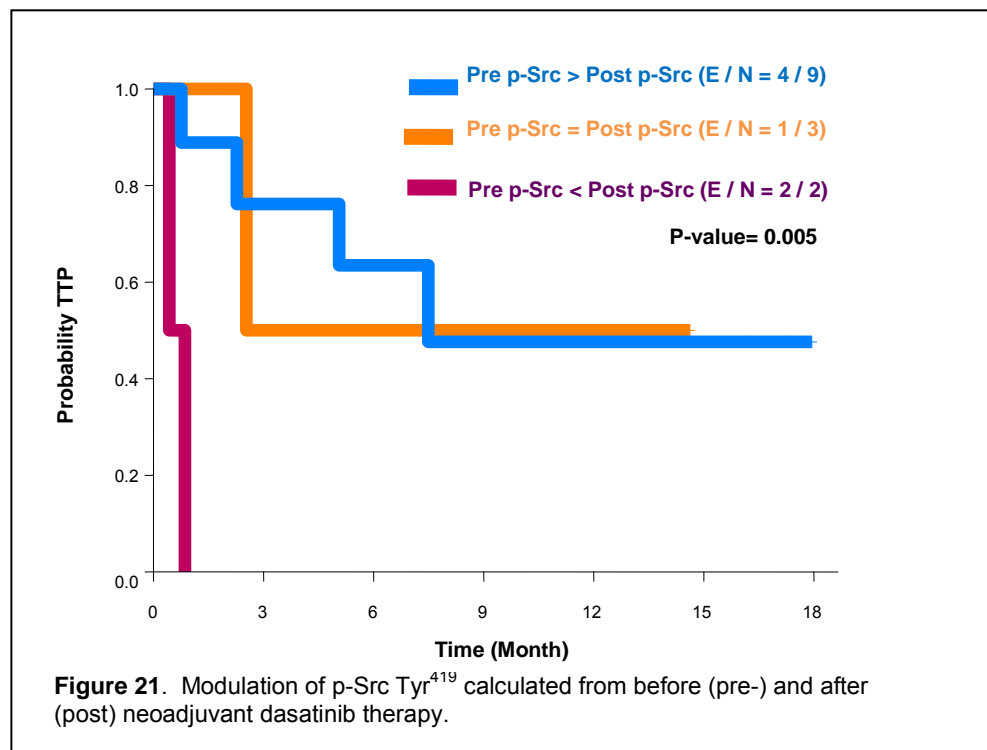
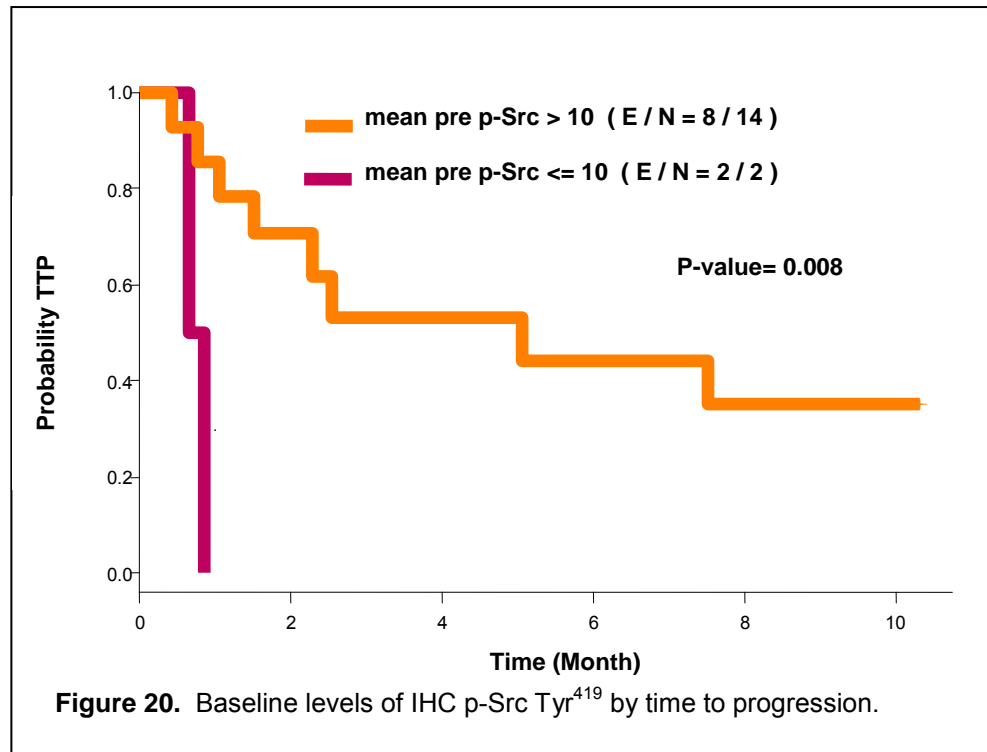
Aim 3: Explore and develop new therapeutic targets and treatment strategies for MPM in tumor specimens collected from Specific Aim1 and in MPM cell lines.

- 3a. Determine key signaling pathways involved in tumor resistance or sensitivity to dasatinib using therapeutic target-focused (TTF) molecular and global gene expression profiling on MPM tumor specimens pre- and post- induction dasatinib therapy.
- 3b. Determine the sensitivity of a panel of MPM cell lines to targeted agents tested in Project 1 via TTF profiling and DATs (drug and therapeutic target siRNA).

Summary of Research Findings

A total of 35 patients have been enrolled on the trial from April 2008 to May 2010. Of these 35 patients, 25 have completed treatment and 17 have successfully completed ESS, neoadjuvant dasatinib, and P/D (n=10) or EPP (n=7). (*SD: Waiting on updates*) The commonly noted side effects of neoadjuvant dasatinib were grade 1-2: anemia, nausea, vomiting, anorexia, fatigue, and anxiety. Grade 3 toxicities included fluid retention, infection (pneumonia), and hypoxia. There are no grade 4-5 toxicities. Post-surgical grade 3 toxicity included anemia, arrhythmia, HTN, and pleural effusion; 1 grade 4 episode of hyperglycemia. To date, neoadjuvant dasatinib does not add significant toxicity to MPM surgical resection.

After 4 weeks of neoadjuvant oral dasatinib therapy, we noted 1 P/D, 13 S/D, and 2 minor responses, and 1 patient that did not receive the second PET-CT. In the initial analysis of IHC Src Tyr419 in 13 patients (**Figure 20**), higher baseline levels of p-Src Tyr419 predicted for an improved TTP with dasatinib therapy (p=0.008). Also, patients who had significant modulation of p-Src Tyr419 after dasatinib therapy had improved TTP (p=0.005) (**Figure 21**). In addition, peripheral mononuclear blood cells (PMBC) and serum samples collected during the surgical resection from 36 MPM patients are available. We have also collected, banked, and characterized MPM tumor tissue from 10 patients enrolled in the clinical trial who underwent video-assisted thoracoscopy (VAT) and extrapleural pneumonectomy (EPP). In collaboration with the Pathology Core, a MPM tissue microarray (TMA) was constructed containing 76 surgically resected tumor cases, including epitheloid, sarcomatoid and biphasic histology types, with well-annotated clinicopathologic information.



The MPM TMAs have been utilized to characterize the expression of several markers, including markers related to epithelial-to-mesenchymal transition (EMT; 5 IHC markers), angiogenesis (PFGFR β ; 2 IHC markers and FISH for the gene), and cell membrane transporters (5 IHC markers). In addition, MPM cell lines acquired from Dr. Harvey Pass are currently being characterized by IHC using 7 different markers to distinguish them as authentic mesothelioma cell lines, including cytokeratin 5/6, calretinin, mesothelin, CEA, B72.3, CD15, and TTF-1. Finally, the expression of total Src and p-Src (Tyr 416), as well as Ki67, was examined by IHC in nearly 100 MPM tissue samples obtained from patients enrolled in the dasatinib clinical trial. Analysis of these data is continuing.

Key Research Accomplishments

- Demonstrated that there is a subpopulation of MPM patients that may derive clinical benefit from oral dasatinib therapy.
- MPM is a very heterogeneous tumor. Molecular profiling will be necessary to ultimately optimize targeted therapy in this disease.
- Illustrated that higher baseline levels of p-Src Tyr⁴¹⁹ is predictive for improved PFS with dasatinib.
- Suggested that modulation of p-SrcTyr⁴¹⁹ is a reasonable pharmacodynamic marker for dasatinib treatment.
- Proven that the infrastructure from this novel clinical is feasible and has the potential to advance the field of MPM in personalized medicine.
- Collected, banked, and characterized MPM tumor tissue from 10 patients enrolled in the clinical trial who underwent video-assisted thoracoscopy (VAT) and extrapleural pneumonectomy (EPP).
- Constructed a MPM tissue microarray (TMA) containing 76 surgically resected tumor cases.
- Examined the expression of total Src and p-Src (Tyr 416), as well as Ki67, by IHC in nearly 100 MPM tissue samples obtained from patients enrolled in the dasatinib clinical trial.

Conclusions

There is preliminary evidence that a subgroup of MPM patients gain clinical benefit from dasatinib therapy and that baseline p-Src Tyr419 levels in MPM tumor tissue may be predictive of TTP. This is the first targeted therapy neoadjuvant trial to potentially identify a predictive biomarker in MPM.

Project 5: Development of a Novel Multi-Biomarker System Using Quantum Dot Technology for Assessments of Prognosis of NSCLC and Prediction of Outcome of EGFR-Targeted Therapy

(Leader: Dr. Zhuo (Georgia) Chen; Co-Leaders: Drs. Fadl o Khuri, Dong Shin, Ruth O'Regan, Shi-Yong Sun)

Quantum dots (QDs) provide sharper fluorescent signals than organic dyes and can detect multi-biomarkers simultaneously in the same material, allowing quantification and correlation of molecular signature with cellular response to targeted therapies.

Hypothesis:

A multi-biomarker system using quantum dot (QD) technology will enhance accuracy in assessment of prognosis of non-small cell lung cancer (NSCLC) and prediction of outcome of epidermal growth factor receptor (EGFR)-targeted therapy.

Specific Aims:

Specific Aim 1: Development of QD-Abs and imaging systems for detection and quantification of multi-biomarkers (MBM) using lung cancer cell lines.

Summary of Research Findings

This aim was completed as reported in the previous annual report.

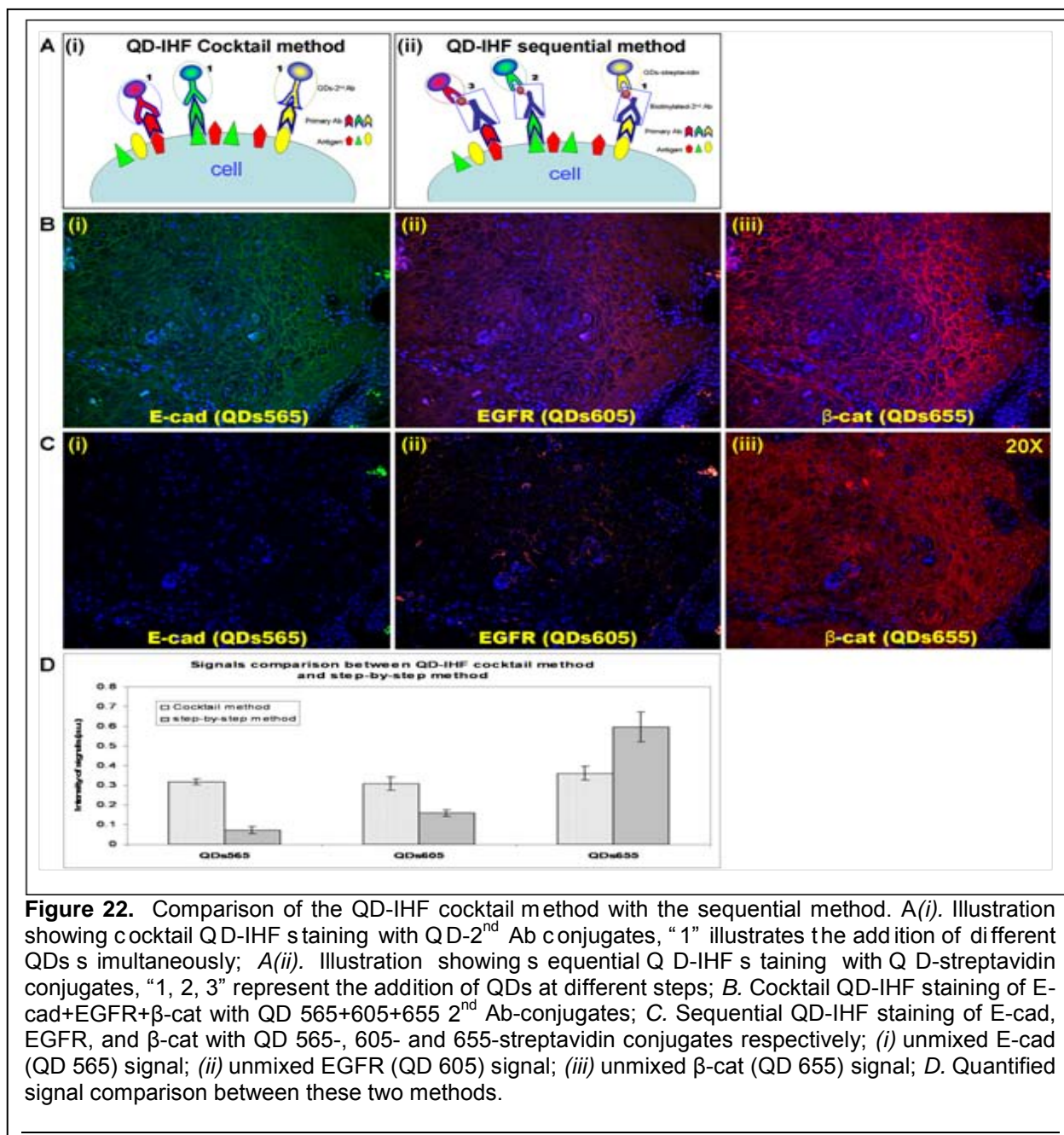
Specific Aim 2: Verification of QD-Abs for detection and quantification of MBM by comparison with conventional IHC using paraffin-embedded tissues and evaluation of their prognostic value in NSCLC.

Summary of Research Findings

Nanoparticle QDs are ideal materials for multiplexed biomarker detection, localization, and quantification. We have optimized the working conditions for the application of QD-based immunohistofluorescence (QD-IHF) in staining of formalin-fixed and paraffin-embedded (FFPE). Our results demonstrated that the QD signal for each multiplexed biomarker was more consistent and stable when using the cocktail method than the sequential method, providing a unique tool for potential research and clinical applications (**Figure 22**).

We have verified QD-IHF methodology, which included: (1) the comparison of single biomarker detection using conventional immunohistochemistry (IHC) with QD-IHF; and (2) the comparison of biomarker signals from samples stained with single QD-IHF in serial sections to biomarker signals from the same proteins but from samples stained simultaneously with multiple QD-IHF. Our study showed that the results obtained from QD-IHF and IHC or single and multiple staining are statistically correlated.

To achieve this specific aim, tissue samples including tumor and adjacent normal from 94 cases of NSCLC with relevant clinical information have been collected. Both IHC and QD-IHF stains of these tissues have been completed, and statistical analysis of QD-IHF is still ongoing. For quantification of IHC results, Weighted Index {WI = [percentage of positive stain x intensity score (0, 1+, 2+, and 3+)] x 100} was recorded. QD-IHF signals were quantified as percentage of co-localization (Co%) of EGFR, E-cadherin, and β -catenin. Biological evidence has shown that activation of EGFR by EGF induces internalization and enhances degradation of E-cadherin. This process reduces co-localization of EGFR with E-cadherin and β -catenin, resulting in low Co%. Therefore, using QD-IHF in detection of multiplex biomarkers on the same tissue will reflect biological event more accurately than single biomarker detection. Preliminary statistical analysis showed that the adjacent normal tissues had significantly higher Co% than tumor tissues from the same patient ($p = 0.0006$ by paired t-test). Furthermore, patients without a smoking history had significantly higher Co% than those who smoked ($p = 0.001$ by Logistic Regression). Further statistical analysis is currently in progress.



In addition, we have studied another signaling pathway which may contribute to the resistance to EGFR targeting. Some research has shown that cancer cells undergoing epithelial mesenchymal transition (EMT) acquire resistance to EGFR-TKIs, though the underlying mechanism remains unknown. Our preliminary study in NSCLC cell lines showed that the G-protein coupled chemokine receptor CXCR4 could activate both p-AKT and p-ERK pathways, which have been reported to induce EMT. In this study, we hypothesized that CXCR4 may be an alternative route for cancer cells to bypass EGFR to activate the downstream pathway under an EMT phenotype. Initially, IHC was used to detect the expressions of CXCR4, p-AKT, p-ERK, and E-cadherin in 94 clinical NSCLC samples and the previously described WI was used to

quantify the expression level of these markers. Western blot was used to detect the expression of CXCR4 and activation of its downstream pathways in NSCLC cell lines. Our result showed that CXCR4 was detected in 98.9% (93/94) of lung cancer samples; however, of greater interest were the three distinctive expression patterns for CXCR4 including cytoplasmic, nuclear, and cytoplasmic/nuclear expression (**Figure 23**). Since it seems unreasonable to see a transmembrane receptor located in nuclei, we confirmed nuclear localization of CXCR4 by Western blot using nuclear extractions from NSCLC cell lines.

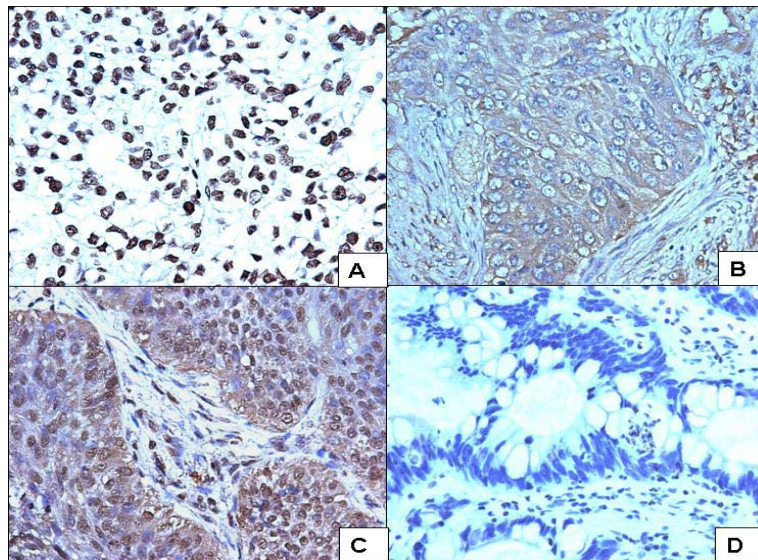


Figure 23. Different staining patterns of CXCR4 in human NSCLC specimens. A. Nuclear staining pattern, the positive signal only located in nuclei (n=22). B. Cytoplasmic staining pattern, the positive signal only located in cytoplasm (n=5). C. Cytoplasmic+nuclear staining pattern, the positive signal distributed both in nuclei and cytoplasm (n=66). D. Negative staining pattern (n=1).

Furthermore, we evaluated the correlation of different locations of CXCR4 with p-AKT, p-ERK, and E-cadherin by ANOVA statistical analysis (**Figure 24**). The result showed the expression level of p-AKT and p-ERK are both significantly higher ($p=0.0024$ and $p=0.0121$, respectively) in CXCR4 cytoplasmic and

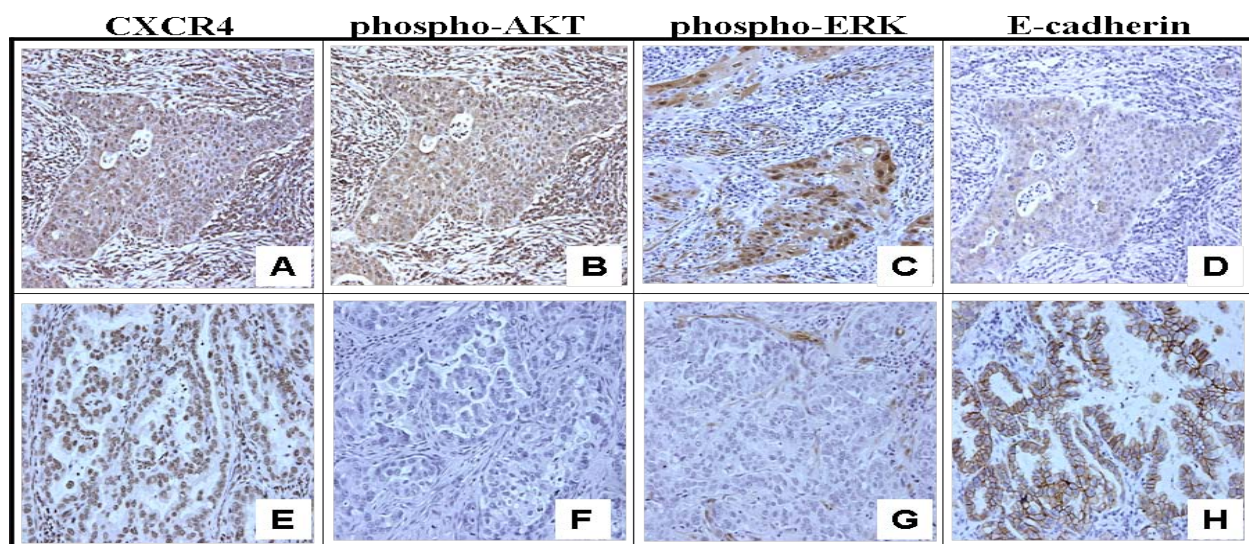


Figure 24. Localization of CXCR4 is correlated to loss of EMT marker and activation of downstream signaling pathways. Two representative cases of squamous cell carcinoma and adenocarcinoma are presented. A, B, C, D were from the same patient tissue in which CXCR4 was located in both cytoplasm and nuclei. Both AKT and ERK1/2 pathways were activated in this patient, while the expression of epithelial marker E-cadherin was dramatically decreased and distributed in a diffuse pattern rather than a normal membrane staining pattern. E, F, G, H were from another patient whose CXCR4 was located only in nuclei. No activation of AKT and ERK1/2 was detected, while E-cadherin was expressed in a high level and well distributed in a linear pattern.

cytoplasmic/nuclear localization groups than those in CXCR4 nuclear localization group. The expression level of E-cadherin in the CXCR4 cytoplasmic and cytoplasmic/nuclear localization groups was also significantly lower ($p=0.0015$) than that in CXCR4 nuclear staining group (**Figure 25**). Our data suggested that cytoplasmic CXCR4 may contribute to activation of AKT and ERK signaling pathways under EMT, while nuclear localization of CXCR4 may be an inactive form. The functional cytoplasmic CXCR4 may contribute to EGFR-TKI resistance by offering an alternative pathway for tumors to maintain progression. A QD-IHF system for this set of biomarkers is under development.

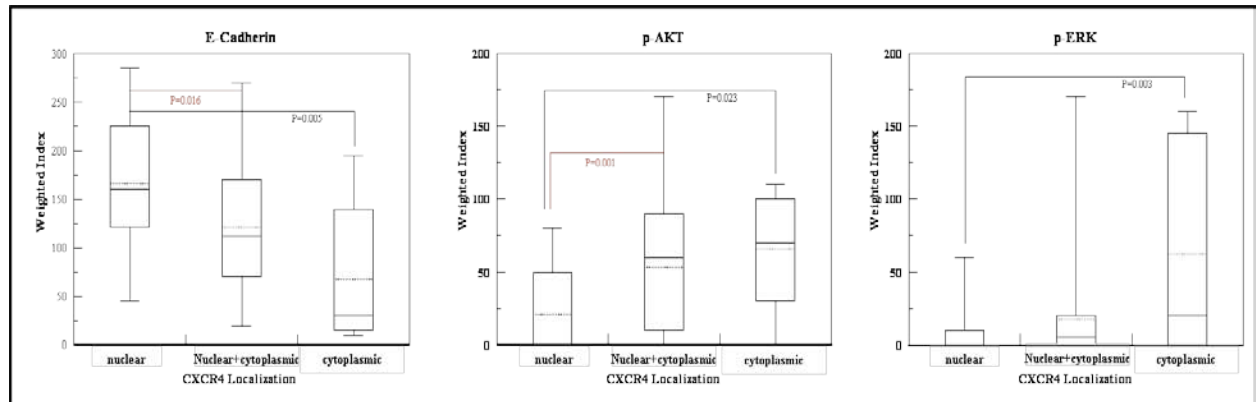


Figure 25. Expression level of p-AKT, p-ERK and E-cadherin are related to the localization of CXCR4. The correlation of different locations of CXCR4 with p-AKT, p-ERK, and E-cadherin was evaluated by two-tail t -test. Significant difference of WI between nuclear and nuclear/cytoplasmic is highlighted in red. Dash-line (----): mean value; Solid-line (—): medium value.

Specific Aim 3: Correlation of the MBM detected by QD-Abs with outcomes of chemotherapies and EGFR- targeted therapy using resectable NSCLC tissues.

Summary of Research Findings

We are currently collecting tissue samples from the patients treated with EGFR-targeted therapy. We may extend this study to head and neck squamous cell carcinoma for further testing and validation of our findings outside the scope of this grant. There is no additional change for this Specific Aim.

Key Research Accomplishments

- Optimized and validated QD-staining conditions for multiplexing three biomarkers, EGFR, E-cadherin, and β -catenin, in both cell lines and FFPE tissues and developed a quantification method for QD signals using the CRi Nuance spectral system.
- Completed staining and image acquisition of the three biomarkers in the 94 pairs of the NSCLC tissues by both IHC and QD-IHF methods.
- Investigated another signal transduction pathway associated with CXCR4. Current study reveals that cellular localization of CXCR4 may correlate to EMT, which has been suggested to be an indicator for resistance to EGFR TKIs. A QD-IHF system for detection and quantification of the relevant biomarkers is under development.

Conclusions

In the past year, we optimized and validated a quantification strategy for using QD-based IHF. These studies were summarized in a publication in *Nano Research* this year, providing a solid foundation for analyzing biomarker expressions in NSCLC tissues. In the past year, we have completed the immunostaining of three biomarkers, EGFR, E-cadherin, and β -catenin, in 94 pairs of patient tissue samples using this strategy. Images of these analyses were acquired and quantification of marker expression has been completed. More statistical analysis will lead to an important answer as to whether quantification of multiplex biomarkers by QD-IHF can provide more accurate correlation to patient's prognosis and the other relevant clinical information than a signal biomarker analysis. In addition, we have identified localization of CXCR4 as a potential indicator for EMT as well as resistance to EGFR targeting. Development of a QD-IHF-based system to quantify CXCR4 localization and its correlation with EMT marker E-cadherin and other relevant biomarkers is the current focus of the project. Building on findings from the first set of biomarkers, this project is moving forward to identify the best marker combination to identify patients who may be resistant to EGFR-targeting therapy.

Pathology Core

(Director: Dr. Ignacio Wistuba)

The Pathology Core is an essential component of the PROSPECT program. The Pathology Core plays an important role by collecting, processing and distributing tissue and serum specimens obtained from clinical trials for non-small cell lung carcinoma (NSCLC) (Project 2) and malignant pleural mesothelioma (MPM; Project 4) for molecular profiles and biomarker analysis.

Our objectives are as follows:

1. Develop and maintain a repository of tissue and serum specimens from patients with NSCLC and MPM.
2. Process NSCLC cell lines and tissue specimens for histopathologic and molecular analyses.
3. Perform and evaluate immunohistochemical (IHC) analysis in human tumor tissue specimens and mouse xenograft tissues.

Objective 1. Develop and maintain repository of tissue and serum specimens from patients with NSCLC and malignant pleural mesothelioma (MPM).

Summary of Research Findings

Selection of lung cancer and mesothelioma specimens available in Thoracic Malignancy Tissue Bank. This goal was achieved during the second year. During the third year the cases selected were processed and distributed for work related to Projects 2, 3 and 4.

Prospective collection and banking of lung cancer and mesothelioma specimens for PROSPECT projects. Since the activation of the PROSPECT project on August 2007, the Pathology Core has collected fresh and formalin-fixed tissue specimens from 444 NSCLC and 31 MPM surgically resected cases (**Table 2**). These cases represent 83% of surgeries performed in our institution during this period. Snap-frozen normal and tumor tissue have been collected in all cases. In addition, we have obtained and banked tumor specimens in RNeasy® (Ambion, Austin, TX) (n=127 samples) consisting of 12% dimethyl sulfoxide (DMSO)-preserved

samples (n=126 samples) and O CT-embedded for frozen sectioning (n=124 samples). Blood specimens (serum and P MBC) were collected from 437 out of 574 surgery patients (76%), processed, and banked. Of interest, both tissue and blood specimens have been obtained in 339 cases (59%). To date, 117 cases of NSCLC treated with neoadjuvant chemotherapy have been collected.

Table 2. Summary of prospectively collected tumor tissue specimens from NSCLC and MPM cases.

Histology	Total Number of Cases	Last Year Number of Cases
Adenocarcinoma	252	92
Squamous cell carcinoma	102	39
Large cell carcinoma	8	4
Other NSCLC	71	30
No tumor present	11	7
Total Lung Tumors	444	172
Malignant Mesothelioma	31	12

Thoracic Malignancy Tissue Bank Database. For tissue collection and banking, an institutional database with bar coding system has been utilized since inception of this grant.

Objective 2. Process NSCLC cell lines and tissue specimens for histopathological and molecular analyses.

Summary of Research Findings

Cell Lines. The establishment of a repository of lung cancer and MPM cell lines was completed during the second year of the grant by the Pathology Core. During the third year, in collaboration with Project 1 (Drs. J. Heymach and J. Minna), we maintained, distributed, and utilized the ~ 70 NSCLC and 2 normal bronchial epithelial cell lines that we have in the repository. In collaboration with Dr. Suraokar, 21 mesothelioma and mesothelial cell lines were DNA-fingerprinted and authenticated as MPM using a panel of 7 immunohistochemistry (IHC) markers (cytokeratin 5/6, calretinin, mesothelin, CEA [carcinoembryonic antigen], B72.3, CD15, and TTF-1). In addition, nucleic acids (DNA and RNA) and protein lysates were extracted from all 21 MPM cell lines and distributed to Project 3 investigators to be molecularly profiled using mRNA Affymetrix, miRNA Agilent, DNA Affymetrix SNP array, and protein RPPA analyses.

Tissue Processing for RNA, DNA and Protein Extractions. As reported last year, tissue processing and nucleic acids (DNA and RNA) extractions of frozen tumor and normal tissue from 613 NSCLCs and 53 MPMs retrieved from the Thoracic Tissue Bank were completed. The processing included a detailed histopathology quality control assessment of tumor and malignant cell content. During the third year, the Pathology Core selected and processed two additional sets of frozen tumor tissue samples for nucleic acid and protein extractions: a) 70 neoadjuvant treated surgically resected NSCLCs; and, b) 23 surgically resected MPMs. Clinicopathological characteristics of all these cases, including treatment and follow-up, were obtained. In addition, the Pathology Core re-processed over 80 NSCLC tumor specimens from the first set of cases for DNA extraction to increase the DNA yield for gene mutation analysis.

Sample Distribution. During the third year, the following samples were distributed by the Pathology Core for molecular profiling to Projects 3 and 4:

RNA. Three aliquots of RNA (1 µg each) from the 249 NSCLC tumors selected by investigators of Project 3 were distributed for miRNA profiling using Agilent platform (1 aliquot; I. Wistuba), and mRNA profiling using Illumina and Affymetrix platforms (2 aliquots; J. Minna). Two aliquots of RNA (1 µg each) from 53 MPM tumors were distributed to Project 3 investigators for miRNA profiling using Agilent platform and mRNA profiling using Affymetrix arrays (I. Wistuba).

DNA. One aliquot of DNA from 162 NSCLC tumors, which are part of the 249 cases selected by investigators of Project 3, were distributed for array Comparative Genomic Hybridization (aCGH) using Agilent 244K platform (1 µg aliquot; I. Wistuba). In addition, two 600 ng aliquots were distributed for gene mutation analysis using sequencing (*KRAS* and *EGFR* mutation analysis) and Sequenom® (20-gene panel) methodologies (I. Wistuba). One aliquot of DNA from 53 MPM tumors was distributed to Project 3 investigators for single nucleotide polymorphism (SNP) Affymetrix array profiling (J. Heymach).

Protein. One frozen tissue specimen from each of the 170 NSCLC tumors (part of the 249 cases selected by Project 3 investigators) was distributed for protein extraction and reverse phase protein array (RPPA) analysis of 176 proteins (J. Heymach's lab). One aliquot of protein lysates from each of the 53 MPM tumors was distributed to Project 3 investigators for RPPA analysis of 176 proteins.

Molecular signatures validation strategies. Project 3 has designed a strategy for gene (mRNA), miRNA, and protein signature validation using the Quantitative Nuclease Protection Assay (qNPA™) for mRNA and miRNA signatures, and immunohistochemistry (IHC) for protein expression. In collaboration with High-Throughput Genomic (HTG), Inc.

(<http://www.htgenomics.com/technology/qnpa>), Project 3 has developed a multiple qNPA™ assays to test the expression of up to 200 genes using both frozen and FFPE tissues from surgically resected NSCLC tumors. Currently, the Pathology Core is processing FFPE tissue samples from all 249-profiled NSCLC cases for qNPA™ analysis.

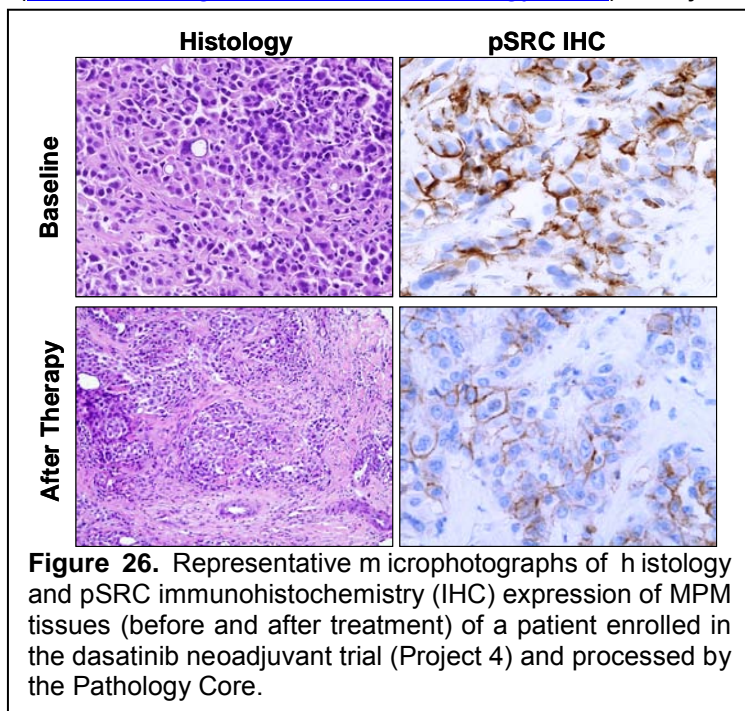


Figure 26. Representative microphotographs of histology and pSRC immunohistochemistry (IHC) expression of MPM tissues (before and after treatment) of a patient enrolled in the dasatinib neoadjuvant trial (Project 4) and processed by the Pathology Core.

Tissue Microarrays (TMAs). An MPM TMA containing 76 tumors was constructed by the Pathology Core during the second year of the grant. During the third year, this TMA was utilized by the Pathology Core to characterize the expression of three key proteins (Mad2a, survivin, and Aurora kinase A) included in the mRNA gene Mitotic Spindle Assembly Checkpoint (MSAC) signature identified to be important in MPM by gene expression profiling in

Project 3 (see objective 3). During the third year, the NSCLC FFPE samples of all 249 tumors selected for molecular profiling by Project 3 investigators have been collected and the histology

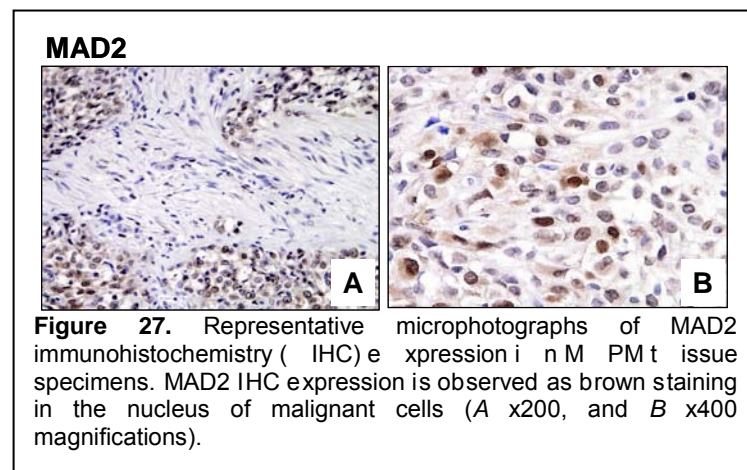
examined by the Pathology Core for TMA construction, which will occur during the fourth year of the grant.

MPM Clinical Trial Tissue Collection and Processing. In collaboration with Project 4 (A. Tsao), the Pathology Core has collected, banked, and characterized MPM tumor tissue from 28 patients enrolled in the dasatinib clinical trial who underwent video-assisted thoracoscopy (VAT) and extrapleural pneumonectomy (EPP). A total of 317 (173 baseline [VAT] and 144 at surgery [EPP]) fresh frozen and formalin-fixed tumor tissue specimens from 28 patients have been obtained, processed, and characterized by the Pathology Core (**Figure 26**).

Objective 3. Perform and evaluate immunohistochemical (IHC) analysis in human tumor tissue specimens and mouse xenograft tumor specimens.

Summary of Research Findings

As previously reported, the Pathology Core has assisted and performed IHC analysis for over 35 markers using TMAs and whole sections in tumor tissue specimens of NSCLC in collaboration with Projects 2 (D. Stewart) and 3 (I. Wistuba). As a result of these analyses, two papers have been accepted for publication (Yuan et al; Solis et al) and one has been recently submitted (Nunez et al).



During the third year, the Pathology Core examined 10 IHC markers using MPM TMA and whole section tissue specimens: a) expression of 3 key proteins, Mad2a, surviving, and Aurora kinase A, included in the mRNA gene MSAC signature identified to be important in MPM by gene expression profiling in Project 3 (**Figure 27**); b) expression of copper transporter receptor 1 (CTR1), glucose 4 (GLUT4) and RHOA, and folate receptor alpha (FOLR1) and reduced folate carrier

1 (RFC1) for Project 2; and, c) expression of total Src and p-Src (Tyr 416) in over 300 MPM tissue samples obtained from patients enrolled in the dasatinib clinical trial for Project 4 (**Figure 26**).

Key Research Accomplishments

- Collected prospective frozen tissue specimens from 165 NSCLC and 12 MPM cases, including 38 NSCLC cases treated with neo-adjuvant chemotherapy.
- Completed the collection of 300 NSCLC tissues originally established for the entire grant period.
- Maintained a NSCLC and MPM cell line repository, and managed the distribution of specimens for molecular profiling in collaboration with PROSPECT Projects 1 and 4.
- Distributed nucleic acids (DNA and RNA) and tissue for protein analyses from large series of NSCLC and MPM tumor specimens with annotated clinicopathological information for molecular profiling.
- Collected, processed, and analyzed over 300 MPM tumor tissue specimens from patients enrolled in the dasatinib clinical trial in collaboration with Project 4.

- Contributed to the publication and submission of 5 manuscripts and 5 abstract presentations in scientific meetings.

Conclusions

During the third grant year, the PROSPECT Pathology Core has achieved and exceeded its goals by prospectively collecting frozen tissue specimens from 165 NSCLC and 12 MPM cases, including 38 NSCLC cases treated with neo-adjuvant chemotherapy. The Pathology Core has played an important role in the processing of NSCLC and MPM tissue and cell line specimens for comprehensive molecular profiling, and in the characterization of tissue specimens on the expression of protein expression by immunohistochemistry.

Biostatistics/Bioinformatics Core

(Director: Dr. J. Jack Lee; Co-Director: Kevin Coombes)

In close collaboration with the Pathology Core and each of the five main projects, the Biostatistics and Data Management Core (BDMC) of the Department of Defense (DoD) PROSPECT Lung Cancer research program is a comprehensive, multi-lateral resource for designing clinical and basic science experiments; developing and applying innovative statistical methodology, data acquisition and management, and statistical analysis; and publishing translational research generated by this research proposal. We deliver planned and tailored statistical analyses for rapid communication of project results among project investigators, and by collaborating with all project investigators to facilitate the timely publication of scientific results.

The main objectives of the Biostatistics and Data Management Core are to:

1. Provide the statistical design, sample size, and power calculations for each project.
2. Develop a secure, internet-driven, Web-based database application to integrate data generated by the five proposed projects and the Pathology Core of the PROSPECT research project.
3. Develop a comprehensive, Web-based database management system for tissue specimen tracking and distribution and for a central repository of all biomarker data.
4. Provide all statistical data analyses, including descriptive analysis, hypothesis testing, estimation, and modeling of prospectively generated data.
5. Provide prospective collection, entry, quality control, and integration of data for the basic science, pre-clinical, and clinical studies in the PROSPECT grant.
6. Provide study monitoring and conduct of the neoadjuvant clinical trial that ensures patient safety by timely reporting of toxicity and interim analysis results to various institutional review boards (IRBs), the U TMDACC data monitoring committee, the DoD, and other regulatory agencies.
7. Generate statistical reports for all projects.
8. Collaborate with all project investigators and assist them in publishing scientific results.
9. Develop and adapt innovative statistical and genomic methods pertinent to biomarker-integrated translational lung cancer studies.

Summary of Research Findings

In the third funding year, the BDMC continued to work with all project investigators in providing biostatistics and data management support. The accomplishments are summarized below.

Biostatistics and Bioinformatics

We have continued to work with clinical investigators in providing the biostatistical support in the development and revision of PROSPECT protocols. We provide statistical report in our monthly project meetings to update the accrual, randomization, demographic data, etc.

We developed and evaluated statistical methodology for comparing various test statistics for response adaptive randomization. We also worked on applying Emax model and use both the interaction index and the bivariate thin plate splines to assess drug interactions in combination studies.

In collaboration with the University of Texas Lung SPORE, we continue to work on developing semantic database models for the kinds of assay data being generated by both PROSPECT projects and the SPORE projects.

We continue to work on developing statistical methods for processing and analyzing the reverse-phase protein array (RPPA) data that continues to be generated as part of the PROSPECT study of lung cancer. We have performed (and continue to perform) analyses of PROSPECT data. Although these analyses have not yet resulted in publications, they are eventually expected to do so. These include:

1. Analyses of immunohistochemically stained tissue microarray data looking at markers of prognosis in lung cancer samples. Univariate analysis identified a number of markers that appear to be related either to important clinical covariates or and/to clinically relevant outcomes (overall survival, disease-free survival, or recurrence-free survival). We have also performed multivariate analyses to identify robust signatures of outcomes, using the same kinds of methods have been developed in the field of gene expression microarrays. These analyses generated three abstracts/posters presented at AACR 2010, one abstract/poster presented at ASCO 2010, two manuscripts submitted, and four manuscripts under preparation.
2. By analyzing paired Illumina and Affymetrix gene expression data from 74 lung cancer cell lines, we have developed a robust gene signature of epithelial-to-mesenchymal transition (EMT; see also Project 1). This signature has been validated in an independent head and neck cancer data set from a third microarray platform, and has recently been validated and shown to predict response to treatment with erlotinib in patients with NSCLC with wild-type EGFR. An abstract describing these findings has been submitted to the European Organization for Research and Treatment of Cancer (EORTC) meeting, and a full manuscript is in preparation.
3. By integrating gene expression data on 74 lung cancer cell lines with DNA methylation data from the Infinium Human Methylation 27 beadchip, we identified coordinated methylation and gene expression patterns that are associated with EMT. These findings suggest that methylation is an important regulator of EMT in lung cancer. An abstract describing these findings has been submitted to the EORTC meeting, and a full manuscript is in preparation.
4. We analyzed an Affymetrix gene expression data set from the mesothelioma samples. Robust clustering identified three genomically-defined subtypes of mesothelioma. These subtypes were only loosely associated with the traditional subclassification by histology. The subtypes were associated with outcome, having statistically significant differences in overall survival. Moreover, at least two of the three subtypes were characterized by high expression of genes in pathways that can potentially be targeted by existing drugs, suggesting possible novel treatments for the disease. A manuscript describing these findings is in preparation.

5. We have begun analyzing a set of Illumina SNP-chip data on lung cancer cell lines in order to assess copy number. Initial results were used to support a manuscript showing that VEGFR2 copy gain is predictive of shorter overall survival in NSCLC patients treated with platinum adjuvant chemotherapy, which has been submitted to JCO. We have also created a web site that allows PROSPECT researchers to browse the results of the preliminary analysis, and continue to analyze the data in order to identify regions of copy number gain or loss that are associated with response to different chemotherapeutic agents.

6. We analyzed a set of microarray data derived from formalin-fixed paraffin-embedded samples from NSCLC patients. We discovered a gene expression signature that predicts prognosis; this signature has been validated using public data available from the director's challenge, and a manuscript describing the findings has been submitted to JCO.

7. Processing and analysis have begun on microRNA array data sets from lung cancer cell lines, lung cancer patient tumors, and mesothelioma patient tumors.

Data Management

PROSPECT Database Development

The PROSPECT database development takes advantage of the revitalization effort from the VITAL program because similar databases were developed in both projects. To tailor for the PROSPECT specific needs, database extensions were made to allow the collection and management of data from multiple studies including the neoadjuvant studies, adjuvant studies, and regular chemotherapy studies. In addition, the PROSPECT database was developed to extend the reVITALization database in VITAL to provide additional clinical, pathological, and biomarker data repositories and tissue tracking. In this funding period, we continued our database development effort and make updates to improve the function and usability of the database.

The SQL Server 2005 database and ASP.NET web application is implemented with VB.net language. Queries and SQL 2005 reports are provided. Secure Socket Layer (SSL) and secured database password are used to keep data transaction protected and confidential. The tissue data includes clinical and pathological data.

1) The database's clinical module contains the following web forms:

- Patient Information
- Social History (Alcohol and Smoking history)
- Medical History
- Other Malignancy
- Treatments (Surgery, Chemotherapy, Radiotherapy and Other Treatments)
- Clinical Staging
- Follow up

2) The pathological module contains the following web forms:

- Primary and Metastasis data (Diagnosis and Surgery Specimens)
- Histology
- Staging and Tumor Information: Cancer staging (TNM classification) is automatically determined by the system based on the tumor information provided.
- Tissue Bank (Frozen Tissue and Paraffin)

3) Reports: Several Excel reports are provided for clinical and pathological module.

1. Clinical Report

2. Pathological Report
3. Patient Report
4. Accession Report
5. General Information Report
6. Other Malignancy Report
7. Surgery Report
8. Chemotherapy Report
9. Radiotherapy Report
10. Other Treatment Report
11. Staging Report
12. Follow up Report
13. Histology Diagnosis Report

4) Dictionaries: The database gives control for the users to update dictionaries; however, dictionary deletion is prohibited.

Key Research Accomplishments

- Continued to provide biostatistics and bioinformatics support for the PROSPECT project.
- Published collaborative papers from data collected in the PROSPECT project.
- Published methodology papers in adaptive clinical trial design, analysis of drug combination studies, and high-throughput genomic data.
- Developed a secured, web-based database application to assist the study conduct.
- Continued with the database maintenance, training and support as followed:
 - Provide data integrity checking and data correction.
 - Database server maintenance and data back-up are performed periodically.
 - Provide training and custom support to end users.
 - Query Generator has been added to allow for users to retrieve specific data of interest quickly. This extensive tool is very user friendly and allows for very complex data selection, sorting and specific field ordering. After a user has created a template within the interface, they have the option of saving this for later data retrieval through the query management options. Once the data has been retrieved, the user can export this to excel for further review and data manipulation.
 - The patient summary report was created with the SQL Server Reporting Services 2005.

Updated screen shots can be found in Appendix below.

Reportable Outcomes:

A web-based database application is developed and deployed at:
https://insidebiostat/Vanguard_CT/login.aspx

Conclusions

In collaboration with clinical investigators, research nurses, Pathology Core, and basic scientists, the Biostatistics and Data Management Core has continued to deliver the biostatistics and data management support as proposed.

KEY RESEARCH ACCOMPLISHMENTS

PROJECT 1

- Completed protein profiling and gene expression profiling for 75 NSCLC, 30 SCLC, and 5 HBEC cell lines.
- Developed an *EGFR* mutation signature that is prognostic in *EGFR* wild-type lung cancer.
- Identified M *ACC1* as highly expressed in lung cancer as an *EGFR* mutant-associated regulator of MET.
- Identified PARP as being highly expressed in SCLC and as a clinically relevant target.
- Developed an EMT gene expression signature that identifies epithelial and mesenchymal groups in NSCLC and HNC cell lines.
- Identified a novel EMT marker, Axl, in NSCLC and HNC.

PROJECT 2

- Collected tumor specimens from 457 lung cancer patients (including 117 who had neoadjuvant chemotherapy).
- Collected blood samples from 461 lung cancer patients (from whom tumor is also available in 345).
- Performed preliminary correlations of several biomarkers with survival and with stage and tumor type, as outlined in more detail in the Project 3 report.
- Performed preliminary assessments of the use of exponential decay nonlinear regression analysis of patient survival curves.
- Defined effect of the demethylating agent decitabine and of time from last therapy on expression, in human tumors, of transporters that may play a role in chemotherapy uptake.
- Identified numerous potential resistance mechanisms that will be investigated further.
- Correlated patient survival and time to recurrence with percent tumor viability post-neoadjuvant chemotherapy.

PROJECT 3

- Completion of comprehensive molecular profiling (mRNA, miRNA, and DNA-aCGH) of a large series (n=209) of chemo-naïve NSCLCs with annotated clinicopathologic information, including outcome after adjuvant chemotherapy.
- Completion of molecular profiling (mRNA, miRNA, and DNA-aCGH) of 53 MPM tissue specimens and 5 cell lines, and identification of molecular signatures associated with outcome of patients with surgically resected MPM and novel MPM potential molecular targets.
- Completion of comprehensive molecular profiling (mRNA, miRNA, and DNA-aCGH) of 40 chemotherapy neoadjuvant-treated NSCLCs with annotated clinicopathologic information, including patients' outcome.
- Characterization of NSCLC tissue specimens for novel biomarkers associated with resistance to chemotherapy in lung cancer, including Nrf2/Keap1 expression, *KDR* copy number, and membrane transporters expression.
- Identification and characterization of phospho-ORM-1 as a novel nicotinic acetylcholine receptor (NACHR)-associated protein and a potential serum marker for lung cancer detection.

PROJECT 4

- Demonstrated that there is a subpopulation of MPM patients that may derive clinical benefit from oral dasatinib therapy.
- MPM is a very heterogeneous tumor. Molecular profiling will be necessary to ultimately optimize targeted therapy in this disease.
- Illustrated that higher baseline levels of p-Src Tyr⁴¹⁹ is predictive for improved PFS with dasatinib.
- Suggested that modulation of p-SrcTyr⁴¹⁹ is a reasonable pharmacodynamic marker for dasatinib treatment.
- Proven that the infrastructure from this novel clinical is feasible and has the potential to advance the field of MPM in personalized medicine.
- Collected, banked, and characterized MPM tumor tissue from 10 patients enrolled in the clinical trial who underwent video-assisted thoracoscopy (VAT) and extrapleural pneumonectomy (EPP).
- Constructed a MPM tissue microarray (TMA) containing 76 surgically resected tumor cases.
- Examined the expression of total Src and p-Src (Tyr 416), as well as Ki67, by IHC in nearly 100 MPM tissue samples obtained from patients enrolled in the dasatinib clinical trial.

PROJECT 5

- Optimized and validated QD-staining conditions for multiplexing three biomarkers, EGFR, E-cadherin, and β -catenin, in both cell lines and FFPE tissues and developed a quantification method for QD signals using the CRI Nuance spectral system.
- Completed staining and image acquisition of the three biomarkers in the 94 pairs of the NSCLC tissues by both IHC and QD-IHF methods.
- Investigated another signal transduction pathway associated with CXCR4. Current study reveals that cellular localization of CXCR4 may correlate to EMT, which has been suggested to be an indicator for resistance to EGFR TKIs. A QD-IHF system for detection and quantification of the relevant biomarkers is under development.

PATHOLOGY CORE

- Collected prospective frozen tissue specimens from 165 NSCLC and 12 MPM cases, including 38 NSCLC cases treated with neo-adjuvant chemotherapy.
- Completed the collection of 300 NSCLC tissues originally established for the entire grant period.
- Maintained a NSCLC and MPM cell line repository, and managed the distribution of specimens for molecular profiling in collaboration with PROSPECT Projects 1 and 4.
- Distributed nucleic acids (DNA and RNA) and tissue for protein analyses from large series of NSCLC and MPM tumor specimens with annotated clinicopathological information for molecular profiling.
- Collected, processed, and analyzed over 300 MPM tumor tissue specimens from patients enrolled in the dasatinib clinical trial in collaboration with Project 4.
- Contributed to the publication and submission of 5 manuscripts and 5 abstracts presentations in scientific meetings.

BIostatistics and Data Management Core

- Continued to provide biostatistics and bioinformatics support for the PROSPECT project.

- Published collaborative papers from data collected in the PROSPECT project.
- Published methodology papers in adaptive clinical trial design, analysis of drug combination studies, and high-throughput genomic data.
- Developed a secured, web-based database application to assist the study conduct.
- Continued with the database maintenance, training and support as followed:
 - Provide data integrity checking and data correction.
 - Database server maintenance and data back up are performed periodically.
 - Provide training and custom support to end users.
 - Query Generator has been added to allow for users to retrieve specific data of interest quickly. This extensive tool is very user friendly and allows for very complex data selection, sorting and specific field ordering. After a user has created a template within the interface, they have the option of saving this for later data retrieval through the query management options. Once the data has been retrieved, the user can export this to excel for further review and data manipulation.
 - The patient summary report was created with the SQL Server Reporting Services 2005.

REPORTABLE OUTCOMES

Publications (Attached in Appendix A)

Byers LA, Sen B, Saigal B, Diao L, Wang J, Nanjundan M, Cascone T, Mills GB, Heymach JV, Johnson FM. Reciprocal regulation of c-Src and STAT3 in non-small cell lung cancer. *Clinical Cancer Research*. 2009 Nov 15;15(22):6852-61. PMID: 19861436.

Gu X, Lee JJ. A simulation study for comparing testing statistics in response-adaptive randomization. *BMC Medical Research Methodology*. 2010 Jun 5;10(1):48. PMID: 20525382.

Huang DH, Peng HX, Su L, Wang DS, Khuri FR, Shin DM, Chen ZG. Optimization and Comparison of Multiplexed Quantum Dot Immunohistofluorescence. *Nano Research*. 3:61-68, 2010.

Kim WY, Perera S, Zhou B, Carretero J, Yeh JJ, Heathcote SA, Jackson AL, Nikolinakos P, Ospina B, Naumov G, Brandstetter KA, Weigman VJ, Zaghlul S, Hayes DN, Padera RF, Heymach JV, Kung AL, Sharpless NE, Kaelin WG Jr, Wong KK. HIF2 α cooperates with RAS to promote lung tumorigenesis in mice. *Journal of Clinical Investigation*. 2009 Aug;119(8):2160-70. PMCID: PMC2719950.

Kong M, Lee JJ. Applying Emax Model and Bivariate Thin Plate Splines to Assess Drug Interactions. *Frontiers in Bioscience*. E2, 279-292, January 1, 2010.

Lee JJ, Lin HY, Liu DD, Kong M. Emax model and interaction index for assessing drug interaction in combination studies. *Frontiers in Bioscience*. E2, 582-601, January 1, 2010.

Solis LM, Behrens C, Dong W, Suraokar M, Ozburn N, Moran C, Corvalan AH, Biswal S, Swisher SG, Bekele BN, Minna JD, Stewart DJ, Wistuba II. Nrf2 and Keap1 abnormalities in non-small cell lung carcinoma and association with clinicopathologic features. *Clinical Cancer Research*. 2010 Jun 9. [Epub ahead of print]. PMID: 20534738.

Takeyama Y, Sato M, Horio M, Hase T, Yoshida K, Yokoyama T, Nakashima H, Hashimoto N, Sekido Y, Gazdar AF, Minna JD, Kondo M, Hasegawa Y. Knockdown of ZEB1, a master epithelial-to-mesenchymal transition (EMT) gene, suppresses anchorage-independent cell growth of lung cancer cells. *Cancer Letters*. 2010 May 6. [Epub ahead of print]. PMID: 20452118.

Tsuta K, Raso MG, Kalhor N, Liu DD, Wistuba II, Moran CA. Histologic features of low- and intermediate-grade neuroendocrine carcinoma (typical and atypical carcinoid tumors) of the lung. *Lung Cancer*. 2010 May 10. [Epub ahead of print]. PMID: 20462655.

Wang J, Wen S, Symmans WF, Pusztai L, Coombes KR. The bimodality index: a criterion for discovering and ranking bimodal signatures from cancer gene expression profiling data. *Cancer Informatics*. 2009 Aug 5;7:199-216. PMCID: PMC2730180.

Xu L, Nilsson MB, Saintigny P, Cascone T, Herynk MH, Du Z, Nikolinakos PG, Yang Y, Prudkin L, Liu D, Lee JJ, Johnson FM, Wong KK, Girard L, Gazdar AF, Minna JD, Kurie JM, Wistuba II, Heymach JV. Epidermal growth factor receptor regulates MET levels and invasiveness through

hypoxia-inducible factor-1alpha in non-small cell lung cancer cells. *Oncogene*. 2010 May 6;29(18):2616-27. PMID: 20154724.

Yuan P, Kadara H, Behrens C, Tang X, Woods D, Solis LM, Huang J, Spinola M, Dong W, Yin G, Fujimoto J, Kim E, Xie Y, Girard L, Moran C, Hong WK, Minna JD, Wistuba II. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS One*. 2010 Feb 9;5(2):e9112. PMCID: PMC2817751.

Manuscripts submitted, in revision, or in review (Attached in Appendix A)

Cascone T, Herynk MH, Du D, Kadara H, Oborn CJ, Nilsson M, Park YY, Lee JS, Ciardiello F, Langley RR, Heymach JV. A role for stromal EGFR activation in resistance to VEGF blockade in human non-small cell lung cancer (NSCLC) xenograft models. Submitted to *Journal of Clinical Investigation*.

Nunez MI, Behrens C, Woods DM, Lin H, Suraokar M, Kadara H, Minna JD, Hofstetter W, Kalhor N, Hong WK, Lee JJ, Franklin W, Stewart DJ, Wistuba II. High expression of folate receptor alpha in lung cancer correlates with adenocarcinoma histology and EGFR mutation. Submitted to *Clinical Cancer Research*.

Xie Y, Xiao G, Coombes KR, Behrens C, Solis L, Raso G, Girard L, Erickson H, Pan X, Roth J, Danenberg K, Minna JD, Wistuba II. Robust gene expression signature from formalin-fixed paraffin-embedded samples predicts prognosis of non-small-cell lung cancer patients. Submitted to *Journal of Clinical Oncology*.

Yang F, Tang X, Behrens C, Varella-Garcia M, Byers LA, Nilsson MB, Lin HY, Wang J, Raso MG, Girard L, Coombes K, Lee JJ, Herbst RS, Minna JD, Heymach JV, Wistuba II. VEGFR2 gene copy gain is predictive of shorter overall survival in nsclc patients treated with platinum adjuvant chemotherapy. Submitted to *Journal of Clinical Oncology*.

Abstracts (Attached in Appendix A)

Byers LA, Wang J, Yordy J, Fan Y-H, Giri U, Shen L, Wistuba I, Girard L, Coombes K, Weinstein J, Minna JD and Heymach JV (2010) Identification of signaling pathways active in small cell lung cancer (SCLC) compared to non-small cell lung cancer (NSCLC) by proteomic profiling. Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#5556.

Cascone T, Herynk M, Du D, Kadara H, Hanrahan EO, Nilsson MB, Lin HY, Lee JJ, Park YY, Lee JS and Heymach JV. Stromal HGF and VEGFR-1 are associated with acquired resistance to VEGFR tyrosine kinase inhibitors in non-small cell lung cancer (NSCLC). 13th World Conference on Lung Cancer, July 2009, San Francisco, CA. Abstract#7956.

Cascone T, Herynk MH, Xu L, Kadara H, Hanrahan EO, Y-H F, Saigal B, Y-Y P, Lee JJ, Langley RR, Jurgensmeier JM, Ryan AJ, Heymach JV. Increased HGF is associated with resistance to VEGFR tyrosine kinase inhibitors (TKIs) in non-small cell lung cancer (NSCLC). Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#376.

Corvalan A, Suraokar M, Zou X, Chow C-W, Gazdar A, Moran CA, Raso G, Mehran R, Tsao A, Wistuba II. Profiling in pharmacologically re-expressed microRNAs by 5-azacytidine and SAHA identified a metastasis associated miR-148b in malignant pleural mesothelioma cell lines. Oral presentation at the 13th World Conference on Lung Cancer, July 2009, San Francisco, CA.

Jayachandran G, Wu X, Coombes K, Wistuba I, Roth J, Ji L. Identification and characterization of phospho-ORM-1 as a novel nicotinic acetylcholine receptor (NACHR)-associated protein and a potential serum marker for lung cancer detection. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Abstract #1624.

Nunez MI, Behrens C, Woods DM, Lin H, Suraokar M, Girard L, Minna JD, Lee JJ, Hofstetter W, Franklin W, Moran CA, Hong WK, Stewart DJ, Wistuba II. Enriched tumor expression of folate transporters correlates with adenocarcinoma histology type, female gender and presence of *EGFR* mutation in non-small cell lung carcinoma. Oral presentation at the 13th World Conference on Lung Cancer, July 2009, San Francisco, CA.

Saintigny P, Byers LA, Zhang L, Yordy J, Tang XM, Girard L, Lang W, Y-H. F, Ji L, Lee JJ, Kim ES, Hong WK, Lippman SM, Herbst RE, Minna JD, Wistuba I, Heymach JV, Mao L. MYC downregulation and chemoresistance in non-small cell lung cancer (NSCLC): Evidence from the Biomarker-Based Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE) program. Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#2172.

Saintigny P, Zhang L, Girard L, Fan Y-H, Lee JJ, Herbst RE, Kim ES, Coombes K, Blumenschein GR, Tsao A, Lam DC, Gerald WL, Beer DG, Tang MX, Lippman SM, Mao L, Hong WK, Wistuba I, Minna JD and Heymach JV (2010) Development and testing of a mRNA expression signature correlated with the presence of *EGFR* mutations in non-small cell lung cancer. AACR-IASCLC Joint Conference: Lung Cancer 2010. Abstract B38.

Stewart DJ, Nunez M, Behrens C, Lin H, Liu D, Lee JJ, Roth J, Heymach J, Wistuba I. High membrane carbonic anhydrase IX expression predicts relapse of resected stage I-II non-small cell lung cancer. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Philadelphia (PA): AACR; 2010. Abstract #4648.

Suraokar M, Nunez M, Chow CW, Mendoza G, Diao L, Moran C, Raso G, Mehran R, Corvalan A, Coombes K, Tsao A, Wistuba II. Deregulation of the mitotic spindle assembly checkpoint pathway in malignant pleural mesothelioma (MPM) tumors revealed by gene expression profiling. Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#2172.

Tsao A, Wistuba I, Mehran R, Rice D, Gil J, Lee JJ, Lin H, Lippman S, Hong WK. Evaluation of Src Tyr419 as a predictive biomarker in a neoadjuvant trial using dasatinib in resectable malignant pleural mesothelioma. *Journal of Clinical Oncology*. 2010 ASCO Annual Meeting Proceedings. Vol 28, No 15S, 2010: 7566.

Wang Y, Huang D, Fu H, Wang D, Koenig L, Khuri FR, Shin DM, Chen Z(G). Cytoplasmic location of CXCR4 is correlated to loss of EMT marker and activation of downstream signaling pathway in non-small cell lung cancer. Proceedings of the 101st Annual Meeting of the

American Association for Cancer Research; 2010 Apr 17-21; Washington, D.C. Philadelphia (PA): AACR; 2010. Abstract#1149.

Yang F, Tang X, Corvalan A, Behrens C, Lin H, Raso G, Lee JJ, Minna JD, Wistuba II. *VEGFR2/KDR* polymorphisms and gene copy gain with outcome in non-small cell lung cancer. Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#374.

Yuan P, Behrens C, Huang J, Spinola M, Prudkin L, Dong W, Yin G, Moran C, Kim E, Zhou B-B, Minna JD, Wistuba II. Expression of stem cell markers in non-small cell lung carcinoma (NSCLC) and correlation with clinico-pathologic features. Oral presentation at the 13th World Conference on Lung Cancer, July 2009, San Francisco, CA.

Yuan P, Kadara H, Behrens C, Tang X, Woods DM, Solis LM, Huang J, Spinola M, Dong W, Yin G, Fujimoto J, Kim E, Xie Y, Girard L, Moran CA, Hong WK, Minna JD, Wistuba II. Sex determining region Y-box 2 is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. Proceedings of the 101st Annual Meeting, American Association of Cancer Research, April 17-21, 2010, Washington, DC. Abstract#5166.

Presentations

Chen, G. Application of Nanotechnology in Cancer Imaging and Treatment. University of Maryland Dental School, November 9, 2009.

Gold KA, Lee JJ, Rice D, Tse W, Stewart D, Wistuba I, Herbst RS, Lippman SM, Hong WK, Kim ES. Phase II pilot study of neoadjuvant docetaxel and cisplatin followed by adjuvant erlotinib in patients with stage I-III non-small cell lung cancer (NSCLC). 2009 ASCO Annual Meeting Proceedings. Vol 27, No 15S (May 20 Supplement), 2009: 7566.

CONCLUSIONS

PROJECT 1: Protein and gene expression profiling of a large number of cell lines has allowed us to identify a prognostic EGFR dependent signature as well as key signaling molecules in NSCLC (MACC1) and SCLC (PARP) that were highly upregulated in those tumor types. Further characterization of these proteins will identify their role in the development of cancer. Also, as in the case with PARP, investigation into the inhibition of these targets may prove to be an effective strategy for the treatment of these diseases. We developed an EMT gene expression signature that classified cell lines into epithelial- or mesenchymal-like group in NSCLC and HNC. This signature was able to predict response to erlotinib in cell lines and *EGFR/KRAS* wild-type patients. Additionally, a novel EMT marker, Axl, was identified through this analysis and may play an important role in the transformation of these tumors as well serve as a potential therapeutic target. These profiles will allow for multiple biomarker analysis and the identification of intracellular signaling pathways that contribute to the sensitivity or resistance to therapeutics. Our findings will be further validated by correlating these gene expression and proteomic profiles of tumor samples with clinical outcomes in samples from the BATTLE-1 trial. Development of predictive markers will assist in guiding treatment selection as well as identify new targets in lung cancer.

PROJECT 2: During this project period, we have completed collection of the planned specimens ahead of schedule and have made good progress on preliminary assessment of these specimens. Molecular profiles will be generated from these specimens and correlated with patient outcomes and resistance to chemotherapy.

PROJECT 3: During the third year, we reached the goal of molecular profiling chemo-naïve surgically resected NSCLCs from patients who received and did not receive adjuvant chemotherapy. Also, we have completed similar profiling in MPM tissue and cell line specimens, devised gene expression signatures associated with MPM patients' outcome, and identified novel potential molecular targets. In addition, we have characterized two novel markers (Nrf2 and *KDR*) associated with response to platinum-based adjuvant chemotherapy in NSCLC.

PROJECT 4: There is preliminary evidence that a subgroup of MPM patients gain clinical benefit from dasatinib therapy and that baseline p-Src Tyr419 levels in MPM tumor tissue may be predictive of TTP. This is the first targeted therapy neoadjuvant trial to potentially identify a predictive biomarker in MPM.

PROJECT 5: In the past year, we optimized and validated a quantification strategy for using QD-based IHF. These studies were summarized in a publication in *Nano Research* this year, providing a solid foundation for analyzing biomarker expressions in NSCLC tissues. In the past year, we have completed the immunostaining of three biomarkers, EGFR, E-cadherin, and β -catenin, in 94 pairs of patient tissue samples using this strategy. Images of these analyses were acquired and quantification of marker expression has been completed. More statistical analysis will lead to an important answer as to whether quantification of multiplex biomarkers by QD-IHF can provide more accurate correlation to patient's prognosis and the other relevant clinical information than a single biomarker analysis. In addition, we have identified localization of CXCR4 as a potential indicator for EMT as well as resistance to EGFR targeting. Development of a QD-IHF-based system to quantify CXCR4 localization and its correlation with EMT marker E-cadherin and other relevant biomarkers is the current focus of the project. Building on findings from the first set of biomarkers, this project is moving forward to identify the best marker combination to identify patients who may be resistant to EGFR-targeting therapy.

PATHOLOGY CORE: During the third grant year, the PROSPECT Pathology Core has achieved and exceeded its goals by prospectively collecting frozen tissue specimens from 165 NSCLC and 12 MPM cases, including 38 NSCLC cases treated with neo-adjuvant chemotherapy. The Pathology Core has played an important role in the processing of NSCLC and MPM tissue and cell line specimens for comprehensive molecular profiling, and in the characterization of tissue specimens on the expression of protein expression by immunohistochemistry.

BIostatISTICS AND DATA MANAGEMENT CORE: In collaboration with clinical investigators, research nurses, Pathology Core, and basic scientists, the Biostatistics and Data Management Core has continued to deliver the biostatistics and data management support as proposed.

References:

1. Stewart D J, Isa JP , Kurzrock R , et al . Decitabine effect on tumor global DNA methylation and other parameters in a phase I trial in refractory solid tumors and lymphomas. Clin Cancer Res 2009;15:3881-8.
2. Stewart D J, Nunez MI, Jelinek J, et al . Decitabine, DNA methylation, time from last therapy and expression of selected membrane transporters in human tumors. Submitted for publication.
3. Stewart DJ. Tumor and host factors that may limit efficacy of chemotherapy in non-small cell and small cell lung cancer. Crit Rev Oncol Hematol.
4. Stewart D J, Behrens C, Roth J, Wistuba II. Exponential Decay Nonlinear Regression Analysis of Patient Survival Curves: Preliminary Assessment in Non-small Cell Lung Cancer. In Press.

APPENDIX A

Abstracts and Publications

APPENDIX B

Biostatistics Workflow and Database Overview

Reciprocal Regulation of c-Src and STAT3 in Non-Small Cell Lung Cancer

Lauren Averett Byers,^{1,2} Banibrata Sen,² Babita Saigal,² Lixia Diao,³ Jing Wang,³ Meera Nanjundan,⁷ Tina Cascone,² Gordon B. Mills,^{4,6} John V. Heymach,^{2,5} and Faye M. Johnson^{2,6}

Abstract Purpose: Signal transducer and activator of transcription-3 (STAT3) is downstream of growth factor and cytokine receptors, and regulates key oncogenic pathways in non-small cell lung cancer (NSCLC). Activation of STAT3 by cellular Src (c-Src) promotes tumor progression. We hypothesized that c-Src inhibition could activate STAT3 by inducing a homeostatic feedback loop, contributing to c-Src inhibitor resistance.

Experimental Design: The effects of c-Src inhibition on total and phosphorylated STAT3 were measured in NSCLC cell lines and in murine xenograft models by Western blotting. c-Src and STAT3 activity as indicated by phosphorylation was determined in 46 human tumors and paired normal lung by reverse phase protein array. Modulation of dasatinib (c-Src inhibitor) cytotoxicity by STAT3 knockdown was measured by MTT, cell cycle, and apoptosis assays.

Results: Depletion of c-Src by small interfering RNA or sustained inhibition by dasatinib increased pSTAT3, which could be blocked by inhibition of JAK. Similarly, *in vivo* pSTAT3 levels initially decreased but were strongly induced after sustained dasatinib treatment. In human tumors, phosphorylation of the autoinhibitory site of c-Src (Y527) correlated with STAT3 phosphorylation ($r = 0.64$; $P = 2.5 \times 10^{-6}$). STAT3 knockdown enhanced the cytotoxicity of dasatinib.

Conclusions: c-Src inhibition leads to JAK-dependent STAT3 activation *in vitro* and *in vivo*. STAT3 knockdown enhances the cytotoxicity of dasatinib, suggesting a compensatory pathway that allows NSCLC survival. Data from human tumors showed a reciprocal regulation of c-Src and STAT3 activation, suggesting that this compensatory pathway functions in human NSCLC. These results provide a rationale for combining c-Src and STAT3 inhibition to improve clinical responses. (Clin Cancer Res 2009;15(22):6852-61)

Lung cancer accounts for 29% of all cancer deaths in the United States, with a 5-year overall survival rate of 15% for all stages (1). Although chemotherapy remains the standard treatment for advanced or metastatic non-small cell lung cancer (NSCLC), response rates do not exceed 35% with frontline therapies and are even lower in the second-line setting (2). Improving our understanding of the signaling pathways that drive tumor behavior is essential for improving clinical outcomes.

One potential therapeutic target in NSCLC for which clinical inhibitors have been developed is cellular Src (c-Src; ref. 3). The Src family consists of nonreceptor tyrosine kinases involved in signal transduction in both normal and cancer cells (4). c-Src is the best characterized and most often involved in cancer progression. c-Src overexpression has been shown in multiple tumor types, in which its activation correlates with shorter survival (reviewed in ref. 3). In NSCLC, c-Src is expressed and

Authors' Affiliations: ¹Division of Cancer Medicine and Departments of ²Thoracic/Head and Neck Medical Oncology, ³Bioinformatics and Computational Biology, ⁴Systems Biology, and ⁵Cancer Biology, The University of Texas M.D. Anderson Cancer Center, and ⁶The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas; and ⁷Department of Cell Biology, Microbiology, and Molecular Biology, University of South Florida, Tampa, Florida
Received 3/29/09; revised 8/6/09; accepted 8/17/09; published OnlineFirst 10/27/09.

Grant support: Physician Scientist Award (to F.M. Johnson); University of Texas Southwestern Medical Center and M.D. Anderson Cancer Center SPORE NIH grant P50 CA070907; Department of Defense grant W81XWH-07-1-0306 01 (PP-1B); Extension of Radiotherapy Research grant P01 CA006294; CCSG functional proteomics core grant P30 CA016672; and

Kleberg Center for Molecular Markers. J.V. Heymach is a Damon Runyon-Lilly Clinical Investigator supported in part by the Damon Runyon Cancer Research Foundation (CI 24-04). L.A. Byers is supported in part by the AACR-AstraZeneca-Prevent Cancer Foundation Fellowship for Translational Lung Cancer Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Faye M. Johnson, Department of Thoracic/Head and Neck Medical Oncology, Unit 432, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030-4009. Phone: 713-792-6363; Fax: 713-792-1220; E-mail: fmjohns@mdanderson.org.

© 2009 American Association for Cancer Research.
doi:10.1158/1078-0432.CCR-09-0767

Translational Relevance

Cellular Src (c-Src) inhibitors represent an exciting new class of targeted drugs that have shown clinical activity in several disease types. However, despite the fact that c-Src overexpression and activation are associated with worse prognosis in non-small cell lung cancer (NSCLC) and c-Src inhibition leads to a universal and profound inhibition of NSCLC invasion, the cytotoxic effects of c-Src inhibition are variable. Understanding the mechanisms of resistance to these drugs is of critical importance so that they can be used most effectively. In these studies, we show that sustained c-Src-inhibition in NSCLC leads to phosphorylation of its downstream target STAT3 and that knockdown of STAT3 increases cytotoxicity induced by c-Src inhibition. These findings support assessment of combinations of drugs that inhibit STAT3, such as JAK kinase inhibitors, with c-Src inhibitors in NSCLC.

activated in both adenocarcinomas and squamous cell carcinomas (5, 6).

c-Src participates in several normal cellular functions during development and adulthood, including cell cycle progression, immune recognition, adhesion, spreading, migration, apoptosis regulation, and differentiation (reviewed in refs. 3, 7). In cancer cells, constitutive activation of c-Src disregulates many of these processes. Inhibition of c-Src activity using both molecular approaches and pharmacologic inhibitors in multiple cancer cell types has been found to lead to reduced anchorage-independent growth (8), decreased proliferation (9), cell cycle arrest (10), decreased tumor growth *in vivo* (11, 12), apoptosis (9), increased susceptibility to anoikis (13), diminished *in vitro* invasion and migration (14, 15), decreased *in vivo* metastasis (12, 16), and decreased *in vivo* vascularity (17). In NSCLC specifically, c-Src inhibition leads to decreased hypoxia-induced vascular endothelial growth factor expression (18). Inhibition of c-Src with a pharmacologic inhibitor (dasatinib) leads to profound and universal *in vitro* inhibition of migration and invasion of NSCLC cells. However, its effect on viability and proliferation is more variable and occurs at concentrations of dasatinib that are near or above the peak plasma concentrations possible in humans (14).

c-Src has multiple downstream substrates that mediate its biological functions in cancer cells. The interaction between c-Src and its substrate focal adhesion kinase (FAK) is essential for normal cell migration and invasion (19). c-Src also regulates downstream proliferation induced by growth factor receptors. Following activation by growth factor receptors, c-Src promotes survival via phosphorylation of the p85 subunit of phosphatidylinositol 3 kinase (PI3K) and thus the AKT pathway, signal transducer and activator of transcription-3 (STAT3), STAT5, and Shc, and thus the Ras/mitogen-activated protein kinase pathway (13, 20, 21). The STAT family of transcription factors, especially STAT3, regulates oncogenic signaling in many different tumor types (22). Indeed STAT3 is required for viral Src-mediated transformation (23). STAT3 can be activated by growth factor receptors or cytokine receptors, usually via non-

receptor tyrosine kinases such as c-Src or janus-activated kinase (JAK) proteins. STAT3 activation leads to the increased expression of downstream targets (e.g., Bcl-XL, cyclin D1, survivin) and increased cell survival, proliferation, and tumor growth *in vivo* (24). Inhibition of STAT3 results in increased apoptosis, decreased proliferation, and decreased tumor size (25, 26). STAT3 activation can also contribute to angiogenesis (27). Hypoxia-induced vascular endothelial growth factor expression is dependent on c-Src activation; this activation of c-Src leads to the downstream activation of STAT3, which binds to the vascular endothelial growth factor promoter with hypoxia-inducible factor-1 α (HIF-1 α).

Although targeting growth factors and signal transduction pathways is a successful strategy in several tumor types, feedback and parallel signaling pathways can limit the efficacy of this approach. Despite c-Src expression in epithelial tumors, including NSCLC, and robust inhibition of c-Src with clinically relevant agents (e.g., dasatinib), the effect of c-Src inhibition on cell survival and proliferation has been modest (14). Defining mechanisms that limit the cytotoxic effects of c-Src inhibitors may result in the development of combinations of therapeutic agents for NSCLC that inhibit metastasis and enhance cytotoxicity. Because c-Src mediates its effects on cancer cell survival and proliferation via diverse substrates, including STATs, in this study we tested our hypothesis that STAT3 may not be inhibited sufficiently by c-Src inhibition in NSCLC to result in clinical effects. We determined that STAT3 was not durably inhibited in NSCLC cell lines and xenografts following c-Src inhibition, making it a candidate pathway for resistance to chemotherapy. Consistent with this result, we found that depletion of STAT3 enhanced the cytotoxicity of c-Src inhibition. In addition, we observed an inverse correlation between c-Src and STAT3 activation levels in untreated primary lung tumors by reverse-phase protein array. The results of these studies support a model in which c-Src and STAT3 are reciprocally regulated in NSCLC tumors, allowing for cancer cell survival following c-Src inhibition.

Materials and Methods

Pharmacologic inhibitors. Dasatinib for *in vitro* studies was provided by Bristol-Myers Squibb and was prepared as a 10 mmol/L stock solution in DMSO. Dasatinib for animal studies was purchased from The University of Texas M. D. Anderson Cancer Center pharmacy. Pyridone 6 was purchased from Calbiochem.

Cell line selection and culture. Human NSCLC cell lines A549 and H226 were obtained from the American Type Culture Collection. H1299, H2009, and H1792 were gifts from Dr. John Minna (Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, Texas). Cell lines with wild-type epidermal growth factor receptors (EGFR) were selected for these studies because EGFR mutations can profoundly affect the response of NSCLC to c-Src inhibition. Three KRAS mutant cell lines were included (A549, H2009, and H1792) because this mutation is common in patients with NSCLC. Cells were grown in monolayer cultures in RPMI 1640 medium containing 10% fetal bovine serum (A549, H226, H1299, and H1792) or RPMI 1640 medium supplemented with hydrocortisone, insulin, transferrin, estradiol, and selenium (HITES) containing 5% fetal bovine serum (H2009) at 37°C in a humidified atmosphere of 95% air and 5% CO₂.

Western blot analysis. Western blot analysis was done to measure protein expression and phosphorylation at 30 min and 7 h after inhibition or specific knockdown of c-Src, STAT3, and/or JAK. Protein levels in treated cells were compared with those in untreated cells at these

same times to control for the effect of confluence or vehicle effects on STAT3 activation. Antibodies used in the Western blot analysis included c-Src and pSTAT3 S727 (Santa Cruz Biotechnology); pSrc Y419, pSTAT3 Y705, STAT3, pFAK Y861, Lyn, Yes, Bcl-XL, survivin, and STAT5 (Cell Signaling Technology); and β -actin (Sigma Chemical Company).

For the Western blot analysis, cells were rinsed and lysed as previously described (28). Equal protein aliquots from cleared lysates were resolved by SDS-PAGE, transferred to nitrocellulose membranes, immunoblotted with primary antibody, and detected with horseradish peroxidase-conjugated secondary antibody (Bio-Rad Laboratories) and enhanced chemiluminescence reagent (Amersham Biosciences).

Transfection with small interfering RNA. To knock down STAT3, the NSCLC cells were harvested, washed, and suspended at a density of 1 million cells/100 μ L of Nucleofector-V solution (Amaxa Corp.). Small interfering RNA (siRNA; 200 pmol/100 μ L) was added to the cell suspension and electroporated using the U-31 Nucleofector program (Amaxa). Immediately after electroporation, 500 μ L of prewarmed RPMI medium were added to the cuvette, and the cells were transferred to 6-well plates. The medium was changed after 16 h. STAT3 and c-Src siRNA were predesigned by siGenome Smartpool (Dharmacon, Inc., a part of Thermo Fisher Scientific) and obtained from Ambion. Controls included cells that were mock-transfected (i.e., without siRNA) and those transfected with a nontargeting (scrambled) siRNA.

MTT, cell cycle, and apoptosis assays. The MTT assay was used to assess cytotoxicity as previously described (28). For each cell line, eight wells were treated with 0, 1, 2, 4, or 8 μ M dasatinib and the IC₅₀. Cell cycle analysis was done as described previously (29). Briefly, cells were fixed and stained with propidium iodide. DNA content was analyzed by fluorescence-activated cell sorting analysis (Becton Dickinson) using ModFit software (Verity Software House). Apoptosis was measured using terminal deoxynucleotidyl transferase biotin-dUTP nick-end labeling staining using the manufacturer's protocol (APO-BRDU kit, Phoenix Flow Systems) as described previously (30). Briefly, fixed cells were incubated with terminal deoxynucleotidyl transferase and Br-dUTP and subsequently incubated with a fluorescein-labeled anti-Br-dUTP antibody and analyzed by fluorescence-activated cell sorting analysis.

Xenograft nude mouse models. All animal procedures were done in accordance with the policies of M.D. Anderson's Institutional Animal Care and Use Committee. Ten female Swiss *nu/nu* strain, 6-week-old mice were used for each xenograft model. Each athymic nude mouse was injected s.c. with 4 million A549 or H226. When visible tumors had developed, dasatinib was administered by oral gavage at a dose of 20 mg/kg/d for 5 d. The mice were euthanized 2 h after the last dose of dasatinib, tumors were dissected, and the mice were examined for distant metastases. The tumors were homogenized and subjected to Western blotting as previously described (30).

Human NSCLC Tumors. Forty-six paired normal lung and NSCLC tumor samples were obtained from surgical specimens in the M. D. Anderson Cancer Center Thoracic Tissue Bank (Table 1). Of the tumors, 22 were squamous cell carcinomas and 24 were adenocarcinomas. The median age of the patients from whom the tumors had been excised was 67 y (range, 48-81 y); 22 (48%) were women, and 44 (96%) were former or current smokers. Six (13%) patients had stage IA cancer, 14 (30%) stage IB, 2 (4%) stage IIA, 13 (28%) stage IIB, 3 (7%) stage IIIA, 6 (13%) stage IIIB, and 2 (4%) stage IV.

Reverse-phase protein array. Protein lysate was prepared from pellets from tumor tissues as previously described (31). Briefly, lysis buffer [1% Triton X-100, 50 mmol/L HEPES (pH 7.4), 150 mmol/L NaCl, 1.5 mmol/L MgCl₂, 1 mmol/L EGTA, 100 mmol/L NaF, 10 mmol/L NaPPi, 10% glycerol, 1 mmol/L phenylmethylsulfonyl fluoride, 1 mmol/L Na₂VO₄, and 10 μ g/mL aprotinin] was added to the samples, followed by microcentrifugation at 14,000 rpm for 10 min. Clear supernatants were collected, followed by protein quantitation using the BCA reaction kit (Pierce Biotechnology, Inc.). The cell lysate was mixed with SDS sample buffer without bromophenol blue [three parts cell lysate plus one part 4 \times SDS sample buffer, which contained 35% glycerol,

8% SDS, 0.25 mol/L Tris-HCl (pH 6.8)]. Before using the cell lysate, 10% β -mercaptoethanol was added. The samples were boiled for 5 min. Then, the samples (each in duplicate) were serially diluted (1:2-1:16) with SDS sample buffer. To each of the diluted samples an equal amount of 80% glycerol/2 \times PBS solution (8 mL of glycerol mixed with 2 mL of 10 \times PBS without Ca²⁺ and Mg²⁺) was added, after which the diluted samples were transferred to 384-well plates. Reverse-phase protein arrays (RPPA) were produced and analyzed as described (52), with slight modifications. Protein arrays were printed on nitrocellulose-coated glass FAST Slides (Schleicher & Schuell BioScience Inc.) by a GeneTAC G3 arrayer (Genomic Solutions) with 48 200- μ m-diameter pins arranged in a 4 \times 12 format. Forty-eight grids were printed at each slide with each grid containing 24 dots. Protein dots were printed in duplicate with five concentrations. Arrays were produced in batches of 15, and occasional low-quality arrays (e.g., with many spot dropouts) were discarded.

Antibody staining of each array was done using an automated BioGenex autostainer. Briefly, each array was incubated with a specific primary antibody: pSrc Y527, pSTAT3 Y705, pSrc Y416 (which binds the pSrc Y419 activation site in humans), pFAK Y576, p-p130Cas Y249, and pPaxillin Y118; Cell Signaling); c-Src and STAT3 (Upstate). The signal was detected using the catalyzed signal amplification (CSA) system according to the manufacturer's recommended procedure (DakoCytomation California, Inc.). In brief, the RPPA slides were blocked for endogenous peroxidase, avidin, and biotin protein activity with I-block (Applied Biosystems) at room temperature for 15 min. After the blocking procedure, the slides were incubated with primary antibody and secondary antibody and diluted in DAKO antibody diluent with a background-reducing compound at room temperature for 20 min each. The slides were then incubated with streptavidin-biotin complex and biotinyl-tyramide (for amplification) for 15 min each, streptavidin-peroxidase for 15 min, and 3,3'-diaminobenzidine tetrahydrochloride chromogen for 5 min. Between steps, each slide was washed with TBS containing 0.1% Tween-20 (TBST). Spot images were quantified using imaging analysis with an HP Scanjet 8200 scanner (Hewlett Packard) with a 256-shade gray scale at 600 dots per inch.

RPPA data processing and statistical analysis. RPPA data were quantified using a SuperCurve method which detects changes in protein level by Microvigena software (VigeneTech) and an R package developed in-house (32). Briefly, the SuperCurve method generates a common logistic curve by pooling data from all samples on the slide. Individual dilution series numbers for each sample are then mapped onto the SuperCurve for quantification. After quantification, data were logarithm-transformed (base 2) for further processing and analyses. Then median-control normalization was applied on the dataset. The statistical analyses were done using R (version 2.7.0). All samples were done in duplicate, and average values were used for analysis. Two sample *t*-tests were used to compare protein levels between normal and tumor tissue; Pearson correlations were used to analyze the association between protein levels in the samples. In all analyses, *P* \leq 0.05 was considered significant.

Results

c-Src inhibition fails to durably inhibit STAT3 in NSCLC cell lines. Western blot analysis of five NSCLC cell lines showed that c-Src phosphorylation was rapidly (30 minutes) and durably (7 hours) inhibited at a site essential for c-Src activation (pSrc Y419 in human c-Src; Fig. 1A). Total c-Src levels were not changed by dasatinib treatment. In contrast to pSrc Y419, in which inhibition was prolonged, STAT3 activation (as shown by pSTAT3 Y705 levels) was transiently inhibited at 30 minutes in A549 and H226 (0.60 and 0.57 times the control value, respectively), followed by reactivation by 7 hours (1.25 and 1.18 times control). In H2009, H1299, and H1792 activation was

Table 1. Clinicopathologic characteristics of patients included in RPPA analysis of NSCLC tumors

Specimen number	Histology	Age at surgery, y	Gender	Race	Smoking status	T stage	N stage	M stage	Overall stage
700	SQ	79	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
708	SQ	61	F	African American	Never	T ₂	N ₁	M ₀	IIB
719	SQ	76	M	Caucasian	Former	T ₂	N ₀	M ₀	IB
720	SQ	67	F	Caucasian	Current	T ₁	N ₁	M ₀	IIA
739	AD	76	F	Caucasian	Former	T ₂	N ₂	M ₁	IV
746	AD	67	F	Caucasian	Former	T ₄	N ₀	M ₀	IIIB
759	AD	60	M	Caucasian	Current	T ₂	N ₂	M ₀	IIIB
763	AD	60	F	Caucasian	Current	T ₂	N ₂	M ₀	IIIA
764	AD	67	F	Caucasian	Current	T ₂	N ₀	M ₀	IB
767	AD	53	F	African American	Current	T ₁	N ₂	M ₀	IIIA
771	SQ	67	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
773	AD	77	F	Caucasian	Former	T ₄	N ₂	M ₁	IV
774	SQ	67	M	African American	Former	T ₄	N ₂	M ₀	IIIB
781	SQ	80	F	Caucasian	Former	T ₁	N ₀	M ₀	IA
782	AD	67	M	Caucasian	Former	T ₄	N ₀	M ₀	IIIB
786	SQ	81	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
795	AD	64	M	Caucasian	Never	T ₁	N ₀	M ₀	IA
799	SQ	61	F	Caucasian	Current	T ₂	N ₀	M ₀	IB
801	AD	63	F	Caucasian	Current	T ₂	N ₁	M ₀	IIB
803	AD	65	F	African American	Current	T ₂	N ₁	M ₀	IIB
804	SQ	77	M	Caucasian	Former	T ₂	N ₀	M ₀	IB
811	SQ	62	M	Caucasian	Current	T ₂	N ₁	M ₀	IIB
813	SQ	70	M	Caucasian	Current	T ₄	N ₀	M ₀	IIIB
816	AD	73	M	Caucasian	Former	T ₂	N ₀	M ₀	IB
833	AD	72	M	Caucasian	Current	T ₁	N ₀	M ₀	IA
844	SQ	67	M	Caucasian	Former	T ₂	N ₁	M ₀	IB
847	AD	60	F	Caucasian	Current	T ₂	N ₀	M ₀	IB
848	AD	48	F	Caucasian	Current	T ₂	N ₀	M ₀	IB
849	AD	57	M	Caucasian	Current	T ₁	N ₁	M ₀	IIA
852	SQ	67	M	Caucasian	Current	T ₂	N ₀	M ₀	IB
857	SQ	71	M	Caucasian	Current	T ₂	N ₀	M ₀	IB
870	SQ	53	M	Caucasian	Current	T ₂	N ₁	M ₀	IIB
877	AD	68	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
879	SQ	67	M	Caucasian	Current	T ₂	N ₁	M ₀	IIB
883	AD	58	F	Caucasian	Former	T ₂	N ₁	M ₀	IIB
884	AD	70	F	Caucasian	Former	T ₁	N ₀	M ₀	IA
890	AD	68	M	Caucasian	Former	T ₂	N ₀	M ₀	IB
896	AD	49	F	Caucasian	Current	T ₁	N ₀	M ₀	IA
899	AD	65	F	Caucasian	Current	T ₁	N ₀	M ₀	IA
905	SQ	80	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
910	AD	74	F	Caucasian	Former	T ₂	N ₀	M ₀	IB
911	SQ	65	F	Caucasian	Former	T ₂	N ₀	M ₀	IB
913	SQ	61	M	Caucasian	Former	T ₂	N ₂	M ₀	IIIA
920	AD	54	F	Caucasian	Current	T ₂	N ₀	M ₀	IB
930	SQ	66	M	Caucasian	Former	T ₂	N ₁	M ₀	IIB
931	SQ	71	F	Caucasian	Current	T ₄	N ₀	M ₀	IIIB

Abbreviations: AD, adenocarcinoma; SQ, squamous cell carcinoma.

seen beginning at 30 minutes (range, 1.20-1.78 times control) and persisted at 7 hours (1.19-1.63 times control; Fig. 1B).

c-Src inhibition leads to initial STAT3 inhibition then reactivation in vivo. After 1, 5, or 7 hours and after 5 days of dasatinib treatment, tumors from mouse xenografts of NSCLC cells were grossly dissected and examined by immunohistochemistry. Tumors were confirmed to consist primarily of NSCLC cells (>90%) with no distant metastases (data not shown). Protein expression and phosphorylation were then measured by Western blot analysis and compared with control (vehicle-treated

mice sacrificed at 7 hours or 5 days; Fig. 2). As expected, dasatinib treatment resulted in pSrc Y419 inhibition at all times. In A549 xenografts, pFAK inhibition was seen beginning at 5 hours. In contrast, but consistent with the *in vitro* data, pSTAT3 Y705 was inhibited at early times (0.58-fold at 5 hours after treatment) but was strongly induced by 7 hours (3.1-fold). pSTAT3 Y705 levels returned to baseline after 5 days of continuous daily treatment. Unlike the *in vitro* studies, the *in vivo* studies showed that total STAT3 levels were also elevated at 7 hours after treatment but were not significantly different from

control levels. In H226 xenografts, pSTAT3 Y705 was inhibited between 5 and 7 hours (0.32- and 0.44-fold at 5 and 7 hours, respectively), but rose above baseline by 5 days (1.7-fold).

c-Src depletion leads to STAT3 activation. To determine whether STAT3 reactivation was downstream of c-Src and not caused by an off-target effect of dasatinib, we examined the effect of c-Src depletion by siRNA at 72 and 96 hours in the NSCLC cell lines A549 and H226. Cell lines were transfected with either scrambled siRNA or c-Src-specific siRNA. In A549 cells, c-Src knockdown decreased total and activated c-Src levels but increased pSTAT3 Y705 by 6.7- and 7.2-fold at 72 and 96 hours, respectively (Fig. 3). In H226 cells, total and activated c-Src levels were suppressed at both time points; pSTAT Y705 was strongly induced by 72 hours (3.5-fold) and returned near baseline by 96 hours (1.5-fold; Fig. 3). Other Src family members, such as Yes and Lyn, were not affected by c-Src knockdown in either cell line.

STAT3 reactivation is JAK-dependent. To determine whether STAT3 reactivation was mediated by JAK proteins, we tested the effect of JAK inhibition on STAT3 activation. NSCLC cell lines were treated with the JAK inhibitor pyridone 6 alone and in combination with dasatinib (Fig. 1A). JAK inhibition alone had no effect on pSrc Y419 but led to complete inhibition of

pSTAT3 Y705 and its downstream targets survivin and Bcl-XL at 7 hours. When combined with dasatinib, JAK inhibition by pyridone 6 prevented STAT3 phosphorylation.

c-Src and STAT3 are reciprocally regulated in NSCLC patient tumors. To evaluate the relationship between c-Src and STAT3 in clinical samples, we used reverse-phase protein array, due to its sensitivity and the small amount of protein required, to quantify the levels of total and phosphorylated Src and STAT3 as well as downstream targets of Src in paired samples of NSCLC tumors and normal lung tissue. Paired *t*-test of tumor and normal tissue from the same patients showed significantly higher c-Src activity in tumors as illustrated by decreased levels of autophosphorylated and inactive c-Src (pSrc Y527; $P = 1.09 \times 10^{-9}$) in tumor tissue as compared with normal lung (Fig. 4A). In contrast, activated STAT3 (pSTAT3 Y705) levels were significantly lower in tumor tissue ($P = 0.006$; Fig. 4A).

Among the tumor specimens, there was a statistically significant correlation between levels of pSrc Y527 and pSTAT3 Y705, with a Pearson correlation coefficient of 0.32 ($P = 0.03$; Fig. 4B). This correlation was also seen when the ratio of pSrc Y527 to total c-Src in tumors was compared with the ratio of

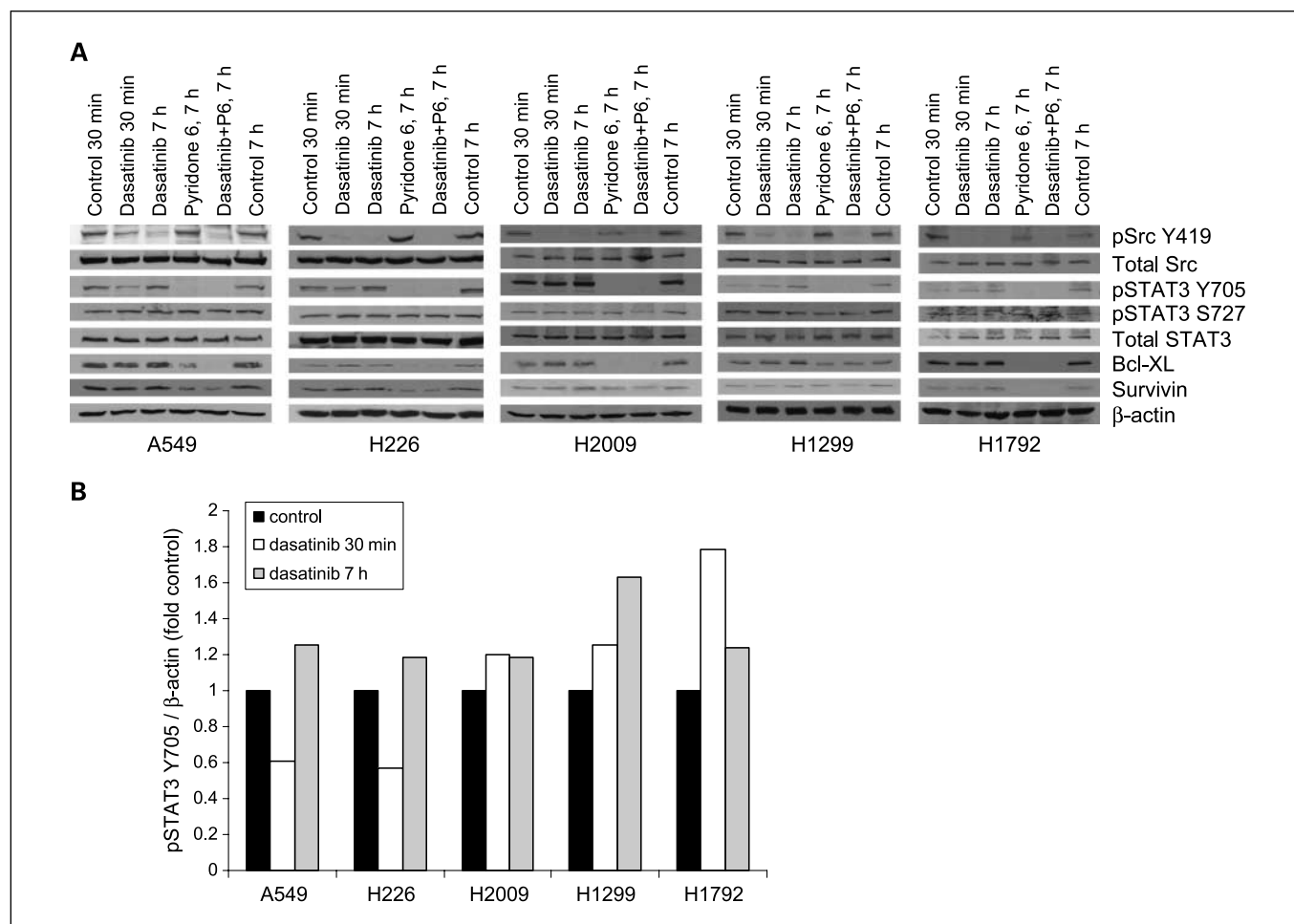


Fig. 1. c-Src inhibition fails to durably inhibit pSTAT3 Y705 *in vitro*. Activation of pSTAT3 Y705 following dasatinib treatment is inhibited by the addition of the JAK inhibitor pyridone 6 (P6). **A**, NSCLC cells were incubated with 100 nmol/L dasatinib, 2.5 μ mol/L pyridone 6, both drugs, or vehicle alone (control) for the indicated times, lysed, and analyzed by Western blotting with the indicated antibodies. **B**, changes in pSTAT3 Y705 were quantified using densitometry. pSTAT3 Y705 levels were normalized to total β -actin and the degree of change was measured between dasatinib-treated cells and control cells corresponding to the same time point (30 min or 7 h).

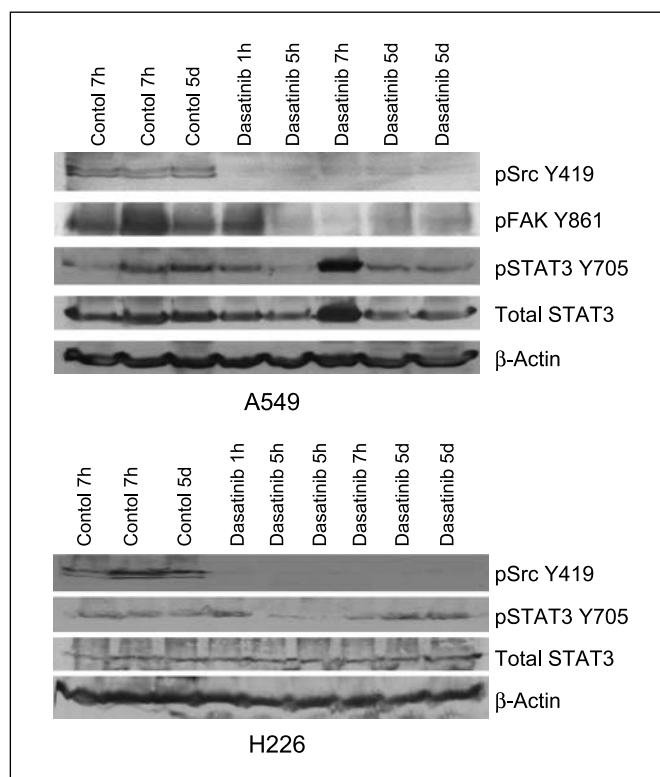


Fig. 2. c-Src inhibition results in STAT3 reactivation *in vivo*. A549 and H226 xenografts were treated daily with dasatinib at 20 mg/kg/d or with vehicle. Tumors were lysed at the times shown and analyzed by Western blotting with the indicated antibodies.

pSTAT3 Y705 to total STAT3 ($r = 0.56$; $P = 5.25 \times 10^{-5}$; Fig. 4B). In addition, the decrease in pSrc Y527 from normal to tumor correlated directly with the decrease in pSTAT3 Y705 from normal to tumor tissue ($r = 0.33$; $P = 0.03$).

Downstream targets of Src (FAK, p130Cas, and paxillin) were also evaluated by RPPA. As expected, activation of these targets (as measured by phosphorylation) was inversely correlated with pSTAT3 Y705 and pSrc Y527. Specifically, correlation coefficients between pSTAT3 Y705 and phosphorylated FAK, p130Cas, and paxillin were $r = -0.50$ ($P = 0.0005$), -0.30 ($P = 0.04$), and -0.42 ($P = 0.004$), respectively (Fig. 4C).

Despite testing several antibodies for RPPA, there is not yet a validated antibody for this assay for the activated form of c-Src (human pSrc Y419). However, using the best performing antibody available, we observed a 1.6-fold higher level of pSrc Y419 in tumors compared with normal tissue ($P = 0.06$). The ratio of pSrc Y419 to total Src was also correlated with the ratio of pSTAT3 Y705 to total STAT3, although this was not statistically significant ($P = 0.67$).

Inhibition of STAT3 with siRNA enhances the cytotoxicity of dasatinib. Because of its role in mediating survival and proliferation, STAT3 induction following c-Src inhibition may represent a mechanism of drug resistance. To examine the biological effects of STAT3 reactivation in NSCLC cells, we evaluated the effect of STAT3 knockdown on cytotoxicity when combined with dasatinib. A549 and H226 cells were transfected with STAT3 siRNA, scrambled (nontargeting) siRNA, or mock-transfected. A549 and H226 cells transfected with STAT3 siRNA showed an 87% and 79% decrease in STAT3 protein levels (re-

spectively) at 48 hours after transfection with STAT3 siRNA as compared with scrambled siRNA (Fig. 5A and B). To determine the biological effect of c-Src inhibition combined with specific depletion of STAT3, cells were treated with dasatinib 48 hours after transfection. A MTT assay was then used to estimate the number of living cells remaining after 72 hours of treatment. Cells with depleted STAT3 were significantly more sensitive to dasatinib than those transfected with scrambled siRNA. In A549 cells, the IC_{50} values were 0.7 and 4 μ mol/L in control and STAT3 siRNA-transfected cells, respectively, for A549 (Fig. 5C), and 5 and 38 μ mol/L for H226 (Fig. 5D). Unlike dasatinib alone, the combination of STAT3 knockdown with dasatinib strongly induced apoptosis in both cell lines (Fig. 5E). Dasatinib alone induced cell cycle arrest in H226 cells, but this was not significantly affected by the addition of STAT3 depletion.

Discussion

In this study, we found that c-Src and STAT3 activation, as shown by phosphorylation status, were reciprocally regulated in NSCLC cell lines, xenografts, and human tumors. Despite an initial inhibition of STAT3 phosphorylation in the A549 and H226 models, prolonged c-Src inhibition resulted in an increase in STAT3 phosphorylation in all NSCLC cell lines tested both *in vitro* and *in vivo*. STAT3 reactivation was JAK-dependent, as illustrated by the observation that reactivation was inhibited by the addition of a JAK inhibitor, pyridone 6. Finally, we showed that inhibition of STAT3 reactivation (by STAT3 siRNA) enhanced the cytotoxicity of dasatinib, showing that this pathway has biological significance. Taken together, these results suggest that STAT3 reactivation may be an important mechanism of resistance to c-Src inhibitors in NSCLC and may be a clinically relevant target for combination therapy.

Previous studies have shown that c-Src is overexpressed in NSCLC and that increased c-Src activity is associated with worse clinical outcome. Clinical investigators are enthusiastic about c-Src inhibitors because specific and potent kinase inhibitors are well tolerated in humans (33). Two such approved anticancer drugs, imatinib and erlotinib, use ATP-competitive kinase

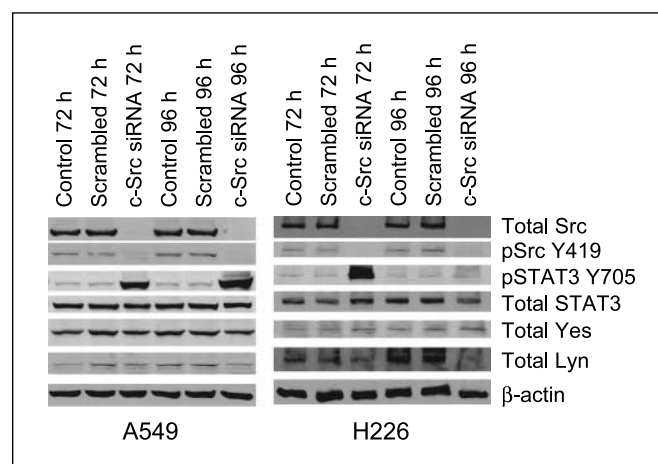


Fig. 3. c-Src depletion results in STAT3 reactivation *in vitro*. A549 and H226 cells were transfected with c-Src-specific siRNA or nontargeting (scrambled) siRNA. Control cells were mock-transfected. Cells were lysed at 72 h or 96 h and analyzed by Western blotting.

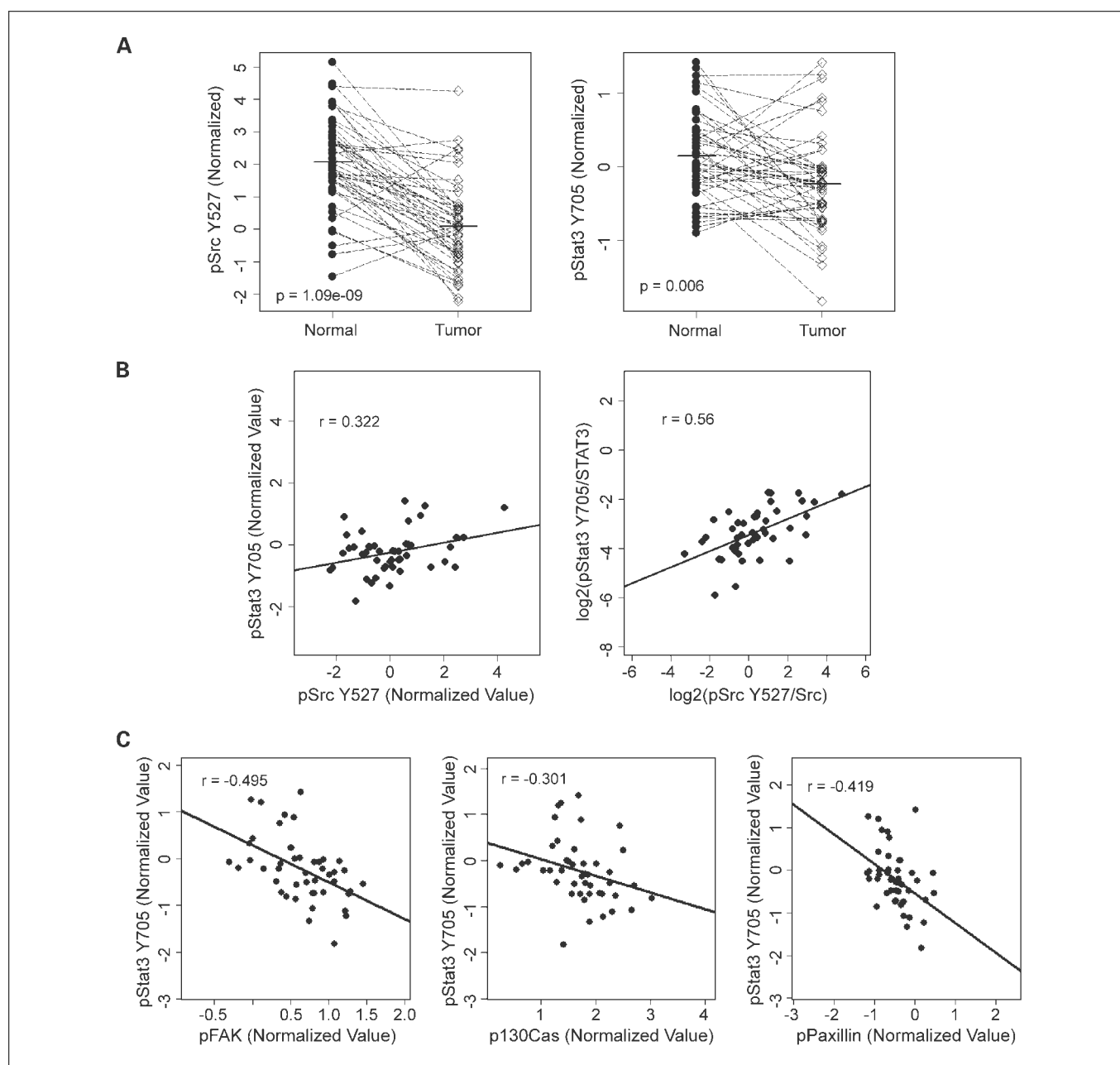


Fig. 4. Phosphorylated c-Src, STAT3, and downstream Src targets were measured in paired normal lung and NSCLC tumor tissues from patients who had previously undergone resection. **A**, tumors had higher mean c-Src activity, as indicated by decreased levels of inactive c-Src (pSrc Y527). Conversely, mean activated STAT3 (pSTAT3 Y705) was lower in tumor tissue than in normal lung. **B**, levels of inactive c-Src (pSrc Y527) correlated directly with activated STAT3 (pSTAT3 Y705) when analyzed as total levels of phosphorylated protein or ratio of phosphorylated protein to total protein. **C**, levels of phosphorylated FAK, p130Cas, and paxillin were inversely correlated with pSTAT3 Y705.

inhibition to inhibit Bcr-Abl (34) and EGFR (35), respectively, proving that kinase inhibition of signal transduction molecules can lead to profound tumor responses. However, thus far c-Src inhibitors have shown limited activity in NSCLC patients. Understanding the mechanisms of resistance to c-Src inhibition in NSCLC will be extremely important for understanding how these drugs can be used more effectively in this disease.

In NSCLC, resistance to tyrosine kinase inhibitors, such as those targeting EGFR, is often due to either activation of the signaling pathway downstream to the drug target (e.g., k-Ras mutations) or signaling through alternate pathways (e.g., c-Met;

refs. 36, 37). Therefore, because STAT3 activity is an important downstream target of c-Src and necessary for c-Src signal transduction, characterization of its relationship to c-Src activity and response to c-Src inhibition was of particular interest. Interesting, these studies did show reactivation of STAT3 in the setting of c-Src inhibition. In cell lines from head and neck squamous cell cancers (30), squamous cell carcinoma of the skin,⁸ and mesothelioma (38), sustained c-Src inhibition also resulted in STAT3 reactivation. This suggests that reciprocal c-Src-STAT3 regulation

⁸ Unpublished data.

exists in multiple tumor types. However, this is the first study to show reciprocal regulation of c-Src and STAT3 in patient tumors.

Three negative feedback loops regulate STAT function after cytokine signaling: SH-2-containing phosphatases, which inactivate JAK by dephosphorylation; protein inhibitors of activated STAT, which are negative regulators of STAT-induced transcription (i.e., downstream of STATs); and suppressors of cytokine signaling, which inhibit JAK activity, facilitate proteosomal degradation of JAK, and compete with STATs for binding to

cytokine receptors (39). Although there are no known positive feedback loops leading to STAT3 activation after its inhibition, loss of a negative feedback loop could play the same role. For example, v-Abl leads to JAK/STAT activation via its disruption of suppressor of cytokine signaling 1 function (SOCS-1) (40).

The concentration (100 nmol/L) of dasatinib was chosen for these studies because it has been shown to completely inhibit c-Src and is relatively specific (41). For example, in intact cells, we observed ~90% reduction in phosphorylated

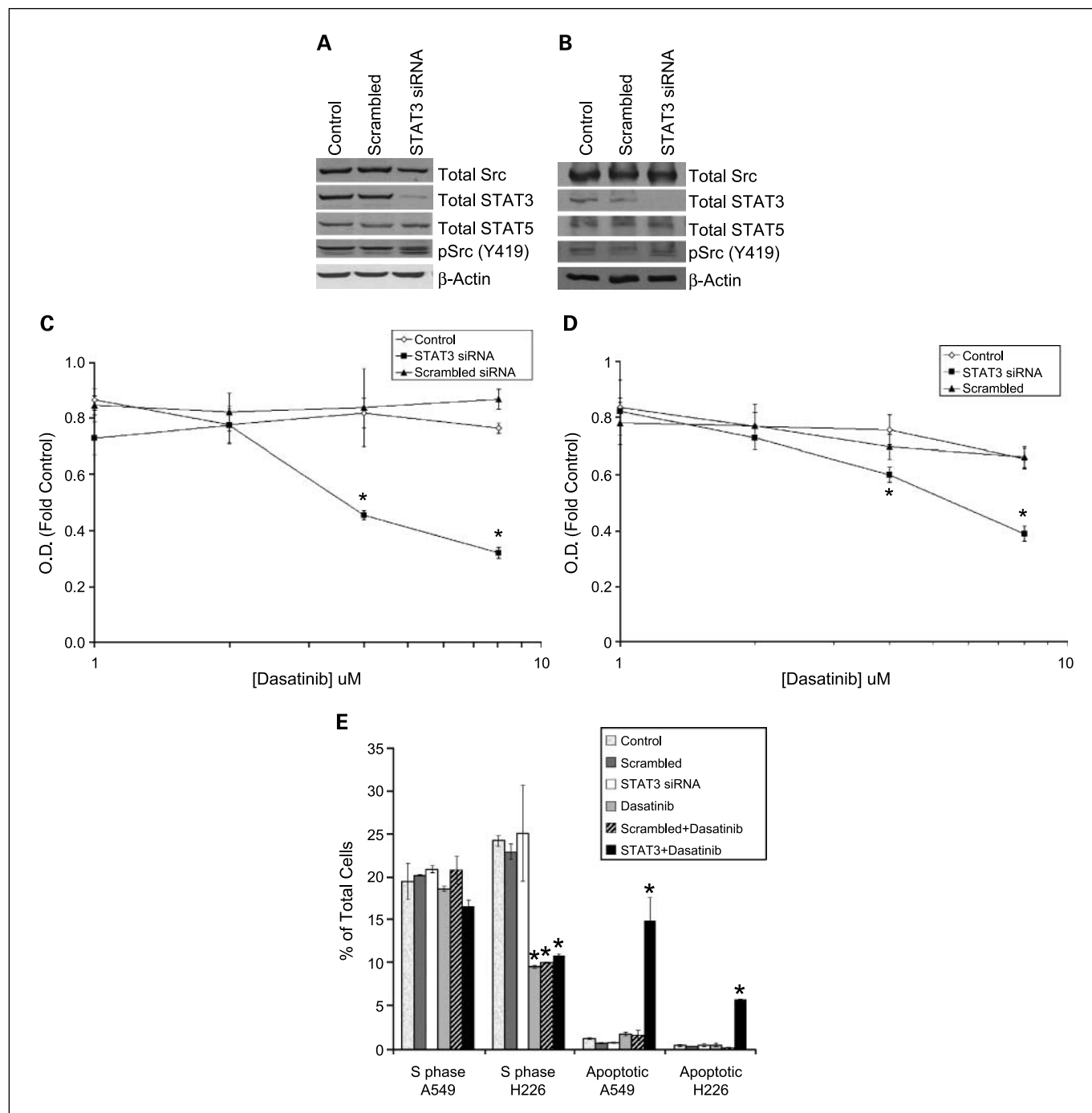


Fig. 5. Depletion of STAT3 enhances cytotoxicity of dasatinib. A549 (A) and H226 (B) cells were transfected with STAT3-specific siRNA, scrambled siRNA, or mock-transfected. Forty-eight hours after transfection, cells were treated with various concentrations of dasatinib. The number of viable A549 (C) and H226 (D) cells after 72 h was evaluated by MTT assay. The percentage of cells in S phase and in apoptosis were then measured under the conditions shown. The control sample was compared with each treatment group and significant differences ($P < 0.05$) are marked with an asterisk (E).

c-Src with >40 nmol/L dasatinib.⁹ In patients treated with dasatinib, plasma levels of approximately 400 nmol/L were reached with sustained levels of 40 to 100 nmol/L (42, 43), supporting the clinical relevance of our selected drug concentration. However, at these concentrations dasatinib also inhibits Abl, PDGFR, Btk, and EphA2 (44, 45). To determine that STAT3 reactivation is downstream of c-Src specifically and not due to an off-target effect of dasatinib, we showed that when c-Src was depleted by siRNA, the levels of pSrc decreased significantly, but pSTAT3 levels increased.

Although we showed an inverse correlation between c-Src and STAT3 activity in human tumors, human NSCLC samples were not available from the post-dasatinib setting to confirm a reactivation of STAT3 in patients following c-Src inhibition. Nevertheless, we believe that our *in vitro* and murine models sufficiently support a rationale for combining c-Src and STAT inhibition in the clinical setting. Three ATP-competitive c-Src inhibitors are being studied in clinical trials: dasatinib,

AZD0530 (AstraZeneca), and SKI-606 (Wyeth). A non-ATP-competitive c-Src inhibitor is also in clinical trial (Kinex pharmaceuticals). Preclinical studies of c-Src inhibitor AZD0530 also showed a reactivation of STAT3 in A549 cells 24 hours after treatment with this drug, further supporting that this may be an important mechanism of resistance across this class of drugs (46). Many JAK inhibitors are being studied in the laboratory and several are in early clinical trial (47–50). Our long-term goal is to use the results of these studies to design clinical trials of these or other more specific c-Src and JAK inhibitors, as available, to improve the survival of patients with NSCLC.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Doris R. Siwak for her contribution to the RPPA experiments.

⁹ Unpublished data.

References

- American Cancer Society. Cancer facts & figures 2008. Atlanta: American Cancer Society; 2008.
- Minna J, Schiller J. In: Kasper DFA, Longo DL, editors. Harrison's principles of internal medicine. 17 ed. New York: McGraw-Hill Medical; 2008.
- Johnson FM, Gallick GE. SRC family nonreceptor tyrosine kinases as molecular targets for cancer therapy. *Anticancer Agents Med Chem* 2007;7:651–9.
- Thomas SM, Brugge JS. Cellular functions regulated by Src family kinases. *Annu Rev Cell Dev Biol* 1997;13:513–609.
- Zhang J, Kalyankrishna S, Wislez M, et al. SRC-family kinases are activated in non-small cell lung cancer and promote the survival of epidermal growth factor receptor-dependent cell lines. *Am J Pathol* 2007;170:366–76.
- Masaki T, Igarashi K, Tokuda M, et al. pp60(c-src) activation in lung adenocarcinoma. *Eur J Cancer* 2003;39:1447–55.
- Yeatman TJ. A renaissance for SRC. *Nat Rev Cancer* 2004;4:470–80.
- Laird AD, Li G, Moss KG, et al. Src family kinase activity is required for signal transducer and activator of transcription 3 and focal adhesion kinase phosphorylation and vascular endothelial growth factor signaling *in vivo* and for anchorage-dependent and -independent growth of human tumor cells. *Mol Cancer Ther* 2003;2:461–9.
- Niu G, Bowman T, Huang M, et al. Roles of activated Src and Stat3 signaling in melanoma tumor cell growth. *Oncogene* 2002;21:7001–10.
- Moasser MM, Srethapakdi M, Sachar KS, Kraker AJ, Rosen N. Inhibition of Src kinases by a selective tyrosine kinase inhibitor causes mitotic arrest. *Cancer Res* 1999;59:6145–52.
- Fleming RY, Ellis LM, Parikh NU, Liu W, Staley CA, Gallick GE. Regulation of vascular endothelial growth factor expression in human colon carcinoma cells by activity of src kinase. *Surgery* 1997;122:501–7.
- Trevino JG, Summy JM, Lesslie DP, et al. Inhibition of SRC expression and activity inhibits tumor progression and metastasis of human pancreatic adenocarcinoma cells in an orthotopic nude mouse model. *Am J Pathol* 2006;168:962–72.
- Windham TC, Parikh NU, Siwak DR, et al. Src activation regulates anoikis in human colon tumor cell lines. *Oncogene* 2002;21:7797–807.
- Johnson FM, Saigal B, Talpaz M, Donato NJ. Dasatinib (BMS-354825) tyrosine kinase inhibitor suppresses invasion and induces cell cycle arrest and apoptosis of head and neck squamous cell carcinoma and non-small cell lung cancer cells. *Clin Cancer Res* 2005;11:6924–32.
- Slack JK, Adams RB, Rovin JD, Bissonette EA, Stoker CE, Parsons JT. Alterations in the focal adhesion kinase/Src signal transduction pathway correlate with increased migratory capacity of prostate carcinoma cells. *Oncogene* 2001;20:1152–63.
- Boyer B, Bourgeois Y, Poupon MF. Src kinase contributes to the metastatic spread of carcinoma cells. *Oncogene* 2002;21:2347–56.
- Ellis LM, Staley CA, Liu W, et al. Down-regulation of vascular endothelial growth factor in a human colon carcinoma cell line transfected with an antisense expression vector specific for c-src. *J Biol Chem* 1998;273:1052–7.
- Sato M, Tanaka T, Maeno T, et al. Inducible expression of endothelial PAS domain protein-1 by hypoxia in human lung adenocarcinoma A549 cells. Role of Src family kinases-dependent pathway. *Am J Respir Cell Mol Biol* 2002;26:127–34.
- Playford MP, Schaller MD. The interplay between Src and integrins in normal and tumor biology. *Oncogene* 2004;23:7928–46.
- Wong BR, Besser D, Kim N, et al. TRANCE, a TNF family member, activates Akt/PKB through a signaling complex involving TRAF6 and c-Src. *Mol Cell* 1999;4:1041–9.
- Arron JR, Vologodskaya M, Wong BR, et al. A positive regulatory role for Cbl family proteins in tumor necrosis factor-related activation-induced cytokine (trance) and CD40L-mediated Akt activation. *J Biol Chem* 2001;276:30011–7.
- Yu H, Jove R. The STATs of cancer-new molecular targets come of age. *Nat Rev Cancer* 2004;4:97–105.
- Yu CL, Meyer DJ, Campbell GS, et al. Enhanced DNA-binding activity of a Stat3-related protein in cells transformed by the Src oncoprotein. *Science* 1995;269:81–3.
- Kijima T, Niwa H, Steinman RA, et al. STAT3 activation abrogates growth factor dependence and contributes to head and neck squamous cell carcinoma tumor growth *in vivo*. *Cell Growth Differ* 2002;13:355–62.
- Song JI, Grandis JR. STAT signaling in head and neck cancer. *Oncogene* 2000;19:2489–95.
- Amann J, Kalyankrishna S, Massion PP, et al. Aberrant epidermal growth factor receptor signaling and enhanced sensitivity to EGFR inhibitors in lung cancer. *Cancer Res* 2005;65:226–35.
- Gray MJ, Zhang J, Ellis LM, et al. HIF-1 α , STAT3, CBP/p300 and Ref-1/APE are components of a transcriptional complex that regulates Src-dependent hypoxia-induced expression of VEGF in pancreatic and prostate carcinomas. *Oncogene* 2005;24:3110–20.
- Johnson FM, Saigal B, Tran H, Donato NJ. Abrogation of signal transducer and activator of transcription 3 reactivation after Src kinase inhibition results in synergistic antitumor effects. *Clin Cancer Res* 2007;13:4233–44.
- Saigal B, Glisson BS, Johnson FM. Dose-dependent and sequence-dependent cytotoxicity of erlotinib and docetaxel in head and neck squamous cell carcinoma. *Anticancer Drugs* 2008;19:465–75.
- Sen B, Saigal B, Parikh N, Gallick G, Johnson FM. Sustained Src inhibition results in signal transducer and activator of transcription 3 (STAT3) activation and cancer cell survival via altered Janus-activated kinase-STAT3 binding. *Cancer Res* 2009;69:1958–65.
- Cheng KW, Lu Y, Mills GB. Assay of Rab25 function in ovarian and breast cancers. *Methods Enzymol* 2005;403:202–15.
- Hu J, He X, Baggerly KA, Coombes KR, Hennessy BT, Mills GB. Non-parametric quantification of protein lysate arrays. *Bioinformatics* 2007;23:1986–94.
- Shah NP. Loss of response to imatinib: mechanisms and management. *Hematology Am Soc Hematol Educ Program* 2005;:183–7.
- Deininger M, Buchdunger E, Druker BJ. The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood* 2005;105:2640–53.

35. Brown ER, Shepherd FA. Erlotinib in the treatment of non-small cell lung cancer. *Expert Rev Anticancer Ther* 2005;5:767-75.
36. Pao W, Wang TY, Riely GJ, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2005;2:e17.
37. Engelman JA, Zejnullahu K, Mitsudomi T, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007;316:1039-43.
38. Tsao AS, He D, Saigal B, et al. Inhibition of c-Src expression and activation in malignant pleural mesothelioma tissues leads to apoptosis, cell cycle arrest, and decreased migration and invasion. *Mol Cancer Ther* 2007;6:1962-72.
39. O'Sullivan LA, Liongue C, Lewis RS, Stephenson SE, Ward AC. Cytokine receptor signaling through the Jak-Stat-Socs pathway in disease. *Mol Immunol* 2007;44:2497-506.
40. Limnander A, Rothman PB. Abl oncogene bypasses normal regulation of Jak/STAT activation. *Cell Cycle* 2004;3:1486-8.
41. Lombardo LJ, Lee FY, Chen P, et al. Discovery of N-(2-Chloro-6-methyl-phenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent anti-tumor activity in preclinical assays. *J Med Chem* 2004;47:6658-61.
42. Morgan J, Demetri G, Wang D, et al. A phase I study of dasatinib, a Src multi-kinase inhibitor, in patients (pts) with GIST and other solid tumors. *Eur J Cancer* 2006;Suppl 4:118.
43. Johnson FM, Chiappori A, Burris H, et al. A phase I study (CA180021-Segment 2) of dasatinib in patients (pts) with advanced solid tumors. *J Clin Oncol* 2007;25:18S:14042.
44. Bantscheff M, Eberhard D, Abraham Y, et al. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* 2007;25:1035-44.
45. Rix U, Hantschel O, Durnberger G, et al. Chemical proteomic profiles of the BCR-ABL inhibitors imatinib, nilotinib, and dasatinib reveal novel kinase and nonkinase targets. *Blood* 2007;110:4055-63.
46. Purnell PR, Mack PC, Tepper CG, et al. The Src inhibitor AZD0530 blocks invasion and may act as a radiosensitizer in lung cancer cells. *J Thorac Oncol* 2009;4:448-54.
47. Pardanani A. JAK2 inhibitor therapy in myeloproliferative disorders: rationale, preclinical studies and ongoing clinical trials. *Leukemia* 2008;22:23-30.
48. Wernig G, Kharas MG, Okabe R, et al. Efficacy of TG101348, a selective JAK2 inhibitor, in treatment of a murine model of JAK2V617F-induced polycythemia vera. *Cancer Cell* 2008;13:311-20.
49. Hexner EO, Serdikoff C, Jan M, et al. Lestaurtinib (CEP701) is a JAK2 inhibitor that suppresses JAK2/STAT5 signaling and the proliferation of primary erythroid cells from patients with myeloproliferative disorders. *Blood* 2008;111:5663-71.
50. Kasper S, Breitenbuecher F, Hoehn Y, et al. The kinase inhibitor LS104 induces apoptosis, enhances cytotoxic effects of chemotherapeutic drugs and is targeting the receptor tyrosine kinase FLT3 in acute myeloid leukemia. *Leuk Res* 2008;32:1698-708.

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

A simulation study for comparing testing statistics in response-adaptive randomization

BMC Medical Research Methodology 2010, **10**:48 doi:10.1186/1471-2288-10-48

Xuemin Gu (xuegu@mdanderson.org)
J. Jack Lee (jjlee@mdanderson.org)

ISSN 1471-2288

Article type Research article

Submission date 4 November 2008

Acceptance date 5 June 2010

Publication date 5 June 2010

Article URL <http://www.biomedcentral.com/1471-2288/10/48>

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

A simulation study for comparing testing statistics in response-adaptive randomization

Xuemin Gu¹ and J. Jack Lee^{2,§}

¹Department of Biostatistics, Division of Quantitative Sciences, The University of Texas M. D. Anderson Cancer Center, P O Box 301402, Unit 1409, Houston, Texas 77230-1402, USA

²Department of Biostatistics, Division of Quantitative Sciences, The University of Texas M. D. Anderson Cancer Center, P O Box 301402, Unit 1411, Houston, Texas 77230-1402, USA

[§]Corresponding author: J. Jack Lee

Email addresses:

XMG: xuegu@mdanderson.org

JJL: jjlee@mdanderson.org

Abstract

Background

Response-adaptive randomizations are able to assign more patients in a comparative clinical trial to the tentatively better treatment. However, due to the adaptation in patient allocation, the samples to be compared are no longer independent. At large sample sizes, many asymptotic properties of test statistics derived for independent sample comparison are still applicable in adaptive randomization provided that the patient allocation ratio converges to an appropriate target asymptotically. However, the small sample properties of commonly used test statistics in response-adaptive randomization are not fully studied.

Methods

Simulations are systematically conducted to characterize the statistical properties of eight test statistics in six response-adaptive randomization methods at six allocation targets with sample sizes ranging from 20 to 200. Since adaptive randomization is usually not recommended for sample size less than 30, the present paper focuses on the case with a sample of 30 to give general recommendations with regard to test statistics for contingency tables in response-adaptive randomization at small sample sizes.

Results

Among all asymptotic test statistics, the Cook's correction to chi-square test (T_{MC}) is the best in attaining the nominal size of hypothesis test. The William's correction to log-likelihood ratio test (T_{ML}) gives slightly inflated type I error and higher power as compared with T_{MC} , but it is more robust against the unbalance in patient allocation. T_{MC} and T_{ML} are usually the two test statistics

with the highest power in different simulation scenarios. When focusing on T_{MC} and T_{ML} , the generalized drop-the-loser urn (GDL) and sequential estimation-adjusted urn (SEU) have the best ability to attain the correct size of hypothesis test respectively. Among all sequential methods that can target different allocation ratios, GDL has the lowest variation and the highest overall power at all allocation ratios. The performance of different adaptive randomization methods and test statistics also depends on allocation targets. At the limiting allocation ratio of drop-the-loser (DL) and randomized play-the-winner (RPW) urn, DL outperforms all other methods including GDL. When comparing the power of test statistics in the same randomization method but at different allocation targets, the powers of log-likelihood-ratio, log-relative-risk, log-odds-ratio, Wald-type Z, and chi-square test statistics are maximized at their corresponding optimal allocation ratios for power. Except for the optimal allocation target for log-relative-risk, the other four optimal targets could assign more patients to the worse arm in some simulation scenarios. Another optimal allocation target, R_{RSIHR} , proposed by Rosenberger and Sriram (*Journal of Statistical Planning and Inference*, 1997) is aimed at minimizing the number of failures at fixed power using Wald-type Z test statistics. Among allocation ratios that always assign more patients to the better treatment, R_{RSIHR} usually has less variation in patient allocation, and the values of variation are consistent across all simulation scenarios. Additionally, the patient allocation at R_{RSIHR} is not too extreme. Therefore, R_{RSIHR} provides a good balance between assigning more patients to the better treatment and maintaining the overall power.

Conclusions

The Cook's correction to chi-square test and Williams' correction to log-likelihood-ratio test are generally recommended for hypothesis test in response-adaptive randomization, especially when

sample sizes are small. The generalized drop-the-loser urn design is the recommended method for its good overall properties. Also recommended is the use of the R_{RSIHR} allocation target.

Background

The response-adaptive randomization (RAR) in clinical trials is a class of flexible ways of assigning treatment to new patients sequentially based on available data. The RAR adjusts the allocation probabilities to reflect the interim results of the trial, thereby allowing patients to benefit from the interim knowledge as it accumulates in the trial. In practice, unequal allocation probabilities are generated based on the current assessment of treatment efficacy, which results in more patients being assigned to the treatment that is putatively superior.

Many RAR designs have been proposed over the years [1-13]. The two key issues extensively investigated are the evaluations of parameter estimations and hypothesis testing. Due to the dependency of assigning new patients based on observed data at that time, conventional estimates of treatment effect are often biased; therefore, efforts have been made to quantify and correct estimation bias [14, 15]. Recent theoretical works have been focused on solving problems encountered in practice, which includes delayed response, implementation for multi-arm trials, and incorporating covariates, etc. [1, 3, 11, 16-18]. Many recent theoretical developments are summarized in [19]. Additionally, in order to compare treatment efficacies through hypothesis testing, studies have been conducted on power comparisons and sample size calculations under the framework of adaptive randomization [20-24]. However, most of the works are based on large sample sizes, and focus on asymptotic properties [4, 12, 22, 25, 26]. But these properties have not been fully studied with small sample sizes. The mathematical challenge imposed by correlated data makes it extremely difficult to derive exact solutions for finite samples. Up to now, only limited results on exact solutions have been available [15, 27], and computer simulation has to be relied upon when sample size is small [23, 24], which is often the case in early phase II trials.

Each RAR design has its own objective, and there are both advantages and disadvantages associated with that objective. It is not our purpose to give a comprehensive assessment of different designs by comparing their advantages and disadvantages. Instead, the primary objective of the present study is to characterize the small sample properties of RAR based on a frequentist approach. In particular, we focus on comparing the performance of commonly used test statistics in RAR of two-arm comparative trials with a binary outcome. Due to the departure from normality caused by data correlation and the discrete nature of a binary outcome, hypothesis tests usually can not be controlled at any given levels of nominal significance. Thus, to make our simulation comparison more relevant, our assessment of hypothesis testing methods and RAR procedures is based on the calculation of both statistical power and the comparison to the nominal type I error rate. Several RAR methods studied in our simulations can assign patients according to a given allocation target, which may be optimal in terms of maximizing the power or minimizing the expected treatment failure. Therefore, we also compare the properties of test statistics at different optimal allocation targets.

The remaining parts of this paper are organized into 4 sections. In the Methods Section, we introduce the adaptive randomization procedures, the optimal allocation rates, and the test statistics used in the simulation. In the Results Section, we present the simulation results. We provide a discussion and final recommendations regarding the RAR methods and hypothesis tests in the Discussion and Conclusions Sections.

Methods

In the present section, we briefly describe the randomization methods, asymptotic hypothesis test statistics, and optimal patient allocation targets that are relevant to our simulations. More detailed information can be found in the corresponding references.

Response-based Adaptive Randomization (RAR)

The RAR procedures investigated in the present study are randomized play-the-winner (RPW) [8] [10], drop-the-loser (DL) [28], sequential maximum likelihood estimation (SMLE) [12], doubly-adaptive biased coin [2, 3], sequential estimation-adjusted urn (SEU) [13], and generalized drop-the-loser (GDL) [11] designs. RPW, DL, SEU and GDL are all urn models in the sense that treatment assignment for each patient can be obtained by sampling balls from an urn. In the usual clinical trial setting, an urn model consists of one urn with different types of balls that represent the different treatments under study. Patients are assigned to treatments by randomly selecting balls from the urn. Initially, the urn contains an equal number of balls for each of the treatment offered in the trial. With the progress of a clinical trial, certain rules are applied to update the contents of the urn in such a way that favors the selection of balls corresponding to the better treatment. For example, under the RPW design, the observation of a successful treatment response leads to the addition of a (>0) balls of the same type to the urn; a lack of success leads to the addition of b (>0) balls of the other type to the urn ($a=b=1$ in our simulation). The limiting allocation rate of patients on treatment 1 is $q_2/(q_1+q_2)$, where $q_1=1-p_1$ and $q_2=1-p_2$ are failure rates, and p_1 and p_2 are success rates (or response rates) for treatments 1 and 2. In the DL model, patients are assigned to a treatment based on the type of ball that is drawn; however a treatment failure results in the removal of a treatment ball from the urn, and

treatment successes are ignored. Due to the finite probabilities of extinction, immigration balls are added to the urn. If an immigration ball is drawn, an additional ball of each type is added. The sampling process is repeated until a treatment ball is drawn. The DL urn design has the same limiting allocation as the RPW urn, but less variability in patient allocation. Both SEU and GDL are urn models allowing fraction number of balls, and can target any allocation rate. For SEU method [13], if the limiting allocation of RPW urn is the target in a two-arm trial, then $\hat{q}_1(i) / [\hat{q}_1(i) + \hat{q}_2(i)]$ balls of type 2 and $\hat{q}_2(i) / [\hat{q}_1(i) + \hat{q}_2(i)]$ balls of type 1 are added to the urn following the allocation of the i th patient. Obviously, the response status of the i th patient is related to the contents of SEU urn only through the calculation of $\hat{q}_1(i)$ and $\hat{q}_2(i)$. For a two-arm GDL urn model [11], when a treatment ball is drawn, a new patient is assigned accordingly, but the ball will not be returned to the urn. Depending on the response of the patient, the conditional average numbers of balls being added back to the urn are b_1 and b_2 for treatments 1 and 2, respectively. Therefore, the conditional average numbers of type 1 and type 2 balls being taken out of the urn can be defined as d_1 and d_2 , where $d_1=1-b_1$ and $d_2=1-b_2$. Immigration balls are also present in a GDL urn. Whenever an immigration ball is drawn, a_1 and a_2 balls are added for treatments 1 and 2, respectively. Zhang et al [11] have shown that the limiting allocation rate of patients on treatment 1 is

$$\frac{n_1}{n} \rightarrow \frac{\frac{a_1}{d_1}}{\frac{a_1}{d_1} + \frac{a_2}{d_2}}. \quad (1)$$

The GDL urn becomes a DL urn when $a_1=1$, $a_2=1$, $b_1=p_1$, and $b_2=p_2$. Although GDL is a general method with different ways of implementation, a convenient approach is taken in our simulation. When a treatment ball is drawn, the ball is not returned, and no ball is added regardless of the

response of the patient. When an immigration ball is drawn, $C\rho_1$ and $C\rho_2$ balls of type 1 and 2 are added, where C is a constant, and ρ_1 and ρ_2 are allocation targets on treatments 1 and 2, which are estimated sequentially using the maximum likelihood estimates (MLE) [11].

The SMLE and doubly-adaptive biased coin design (DBCD) methods can also target any allocation ratios, and SMLE can be implemented as a special case of DBCD method. In DBCD method, the probability of the $(i+1)$ th patient being assigned to treatment 1 is calculated by

$$P_{i+1} = g\left(\frac{n_1(i)}{i}, \rho_1(i)\right), \quad (2)$$

where $r_1 = n_1(i)/i$ and $\rho_1(i)$ are the current allocation rate and estimated allocation rate on treatment 1 [2, 3]. The properties of the DBCD depend largely on the selection of g , which can be considered as a measuring function for the deviation from the allocation target. In the present study, we use the following function suggested by Hu and Zhang [3]:

$$\begin{aligned} g(r, \rho) &= \frac{\rho(\rho/r)^\alpha}{\rho(\rho/r)^\alpha + (1-\rho)[(1-\rho)/(1-r)]^\alpha} \\ g(0, \rho) &= 1 \\ g(1, \rho) &= 0 \end{aligned}, \quad (3)$$

where α is a tuning parameter. When α approaches infinity, the DBCD becomes deterministic and the patients are assigned to the putatively better treatment with probability 1. When α equals to 0, the MLE of ρ becomes the allocation target, and the DBCD method is essentially the same as the SMLE design proposed by Melfi et al [12].

Hypothesis Tests for Two-Arm Comparative Trials

In two-arm comparative trials, the results of a binary outcome variable can be summarized in a 2×2 contingency table (Table 1). The following hypothesis test is often conducted to compare treatment efficacy:

$$\begin{aligned} H_0 : & \quad p_1 = p_2 \\ H_1 : & \quad p_1 \neq p_2 \end{aligned} \quad (4)$$

Nine test statistics for the hypothesis test in (4) are given in Table 2. When relative risk (q_1/q_2) and odds ratio (p_1q_2/q_1p_2) are used to quantify the differences between 2 treatment arms, the test statistics are log-relative-risk and log-odds-ratio, T_{Risk} and T_{Odds} , which are asymptotically distributed as chi-square distribution with one degree of freedom (χ_1^2). When simple difference is used to measure the treatment effect, the applicable test statistics are the Wald-type test statistic T_{Wald} and the score-type test statistics T_{Chisq} , where the variance of simple difference in response rates is evaluated at H_1 or H_0 respectively. Additionally, the test statistics based on the logarithm of likelihood ratio (T_{LLR}) can also be constructed. Besides the 5 commonly used test statistics mentioned above, four modified test statistics are also included in Table 2. T_{MO} is a modified log-odds-ratio test proposed by Gart using the approximation of discrete distributions by their continuous analogues [29]. As shown in Table 2, T_{MO} is essentially a modification to T_{Odds} by adding 0.5 to each cell of a 2×2 table. Similarly, Agresti and Caffo proposed a modification to T_{Wald} by adding 1 to each cell of a contingency table [30], which results in the test statistic T_{MW} in Table 2. T_{MC} is the Cook's continuity correction to chi-square test statistics T_{Chisq} . Williams provided a modification to log-likelihood-ratio test T_{LLR} [31]. The original test statistic T_{LLR} is improved by multiplying a scale factor such that the null distribution of the new test statistic T_{ML} has the same moments as the chi-square distribution.

Since all test statistics in Table 2 are based on χ_1^2 , they are asymptotically equivalent and any one of them can be used for large sample sizes. Meanwhile at small sample sizes, an exact test can be conducted if a model is specified for the data given in Table 1. For example, depending on the

number of fixed margins predetermined for the design, one of the following three models can be applied [32]:

$$\Pr(r_1 | n, n_1, r) = h(r_1 | n, n_1, r), \quad (5)$$

$$\Pr(r_1, r | n, n_1, p) = h(r_1 | n, n_1, r)b(r | n, p), \quad (6)$$

and

$$\begin{aligned} &\Pr(r_1, r, n_1 | n, p, \rho) \\ &= h(r_1 | n, n_1, r)b(r | n, p)b(n_1 | n, \rho), \end{aligned} \quad (7)$$

where $h(r_1 | n, n_1, r)$ represents the hypergeometric distribution of r_1 , $b(r | n, p)$ gives the binomial distribution of r under the null hypothesis of equal response rates ($H_0: p_1 = p_2 = p$), and $b(n_1 | n, \rho)$ denotes the binomial distributions of patients on arm 1 with an allocation ratio of ρ ($\rho = 0.5$ for equal randomization). The p value of exact test can be calculated by maximizing the probability in (5), (6), or (7) over the two nuisance parameters, p and ρ . However, due to data dependency, none of the above three models are directly applicable in adaptive randomization. For example, the allocation ratio ρ in adaptive randomization is a random variable with unknown distribution, and the binomial distribution of n_1 assumed in model (7) is not valid even when the null hypothesis is true. Therefore, in adaptive randomization, unconditional exact tests are not available and asymptotic test statistics such as the ones in Table 2 are required for testing the hypothesis in (4).

Optimal Allocation Ratios

The SMLE, DBCD, SEU, and GDL methods can be utilized to allocate patients based on different allocation targets. The allocation targets simulated in the present study are summarized in Table 3, where R_{Risk} , R_{Odds} , R_{Wald} , R_{Chisq} , and R_{LLR} are optimal allocation ratios maximizing the power of T_{Risk} , T_{Odds} , T_{Wald} , T_{Chisq} , and T_{LLR} respectively, at fixed sample size. The derivation of

T_{Risk} , T_{Odds} , T_{Wald} , T_{Chisq} , and T_{LLR} can be found in [33, 34], which is equivalent to minimizing the variance of corresponding test statistic at a fixed total sample size, and consequently the power of that test statistic is maximized. R_{RSIHR} is a recently proposed allocation target that minimizes the expected total number of failures among all trials with the same power [15, 33]. The general theoretical framework and the practical implementation of optimal allocation in k -arm trials with binary outcomes are discussed and demonstrated by Tymofyeyev et al [35], where the optimization can be conducted over different goals. In practice, the performance of the methodology depends on the chosen RAR procedure. The present simulation study only focuses on two-arm trials, with a goal of maximizing the power or minimizing the total number of failures.

Results

Simulations are conducted at different total numbers of patients ranging from 20 to 200. To simplify the presentation, the results for trials with 30 patients are shown here. When patients are less than 30, adaptive randomization is generally not recommended. For sample size of 100 or larger, all methods yield similar properties in general. For all of the urn models, one ball for each treatment is consistently used as the initial contents of the urn. The number of immigration balls is 1 for both the DL and GDL urns. The tuning parameter of DBCD, α , is fixed at 0 or 2. When α is 0, it results in the SMLE method. The value of the constant C in GDL is 2, which is equivalent to adding 2 treatment balls on average when an immigration ball is drawn. All simulation results are calculated based on 10,000 replicates.

For the purpose of comparison, the true allocation rates are shown in Table 4, and the simulated results for allocation rates on arm 1 are shown in Table 5. Among all RAR methods, DBCD has

the best ability to attain the true allocation target. The comparison between SMLE and DBCD shows that, the allocation becomes more unbalanced and the variation of DBCD decreases with increasing value of tuning exponent α . On the other hand, the patient allocation of SEU results in more balanced mean allocation between two arms with a much larger variation as compared with other RAR methods. The GDL has the lowest variation among the four sequential RAR methods. When R_{RPW} (the same as R_{DL}) is the allocation target, DL urn method has the lowest variation in patient allocation, which is consistent with the fact that the lower bound of the estimate of $\text{Var}(R_{RPW})$ is attained by DL urn [4]. The comparison among allocation targets shows that R_{LLR} has the lowest variation in patient allocation, and the highest variation is usually found at R_{RPW} or R_{Risk} . However, R_{RPW} and R_{Risk} are usually the top two allocation targets that assign more patients to the better treatment. R_{Wald} , R_{Odds} , and R_{LLR} assigns more patients to the worse arm in some simulation cases. Among the three allocation targets that assign more patients to the better treatment (R_{RSIHR} , R_{Risk} and R_{RPW}), R_{RSIHR} has a stable and often the lowest variation in patient allocation.

The simulation results are obtained for five null cases and ten alternative cases, and Table 6 gives the summary by averaging the results over the five null cases and the ten alternative cases for a given RAR method and at a given allocation target. Detailed simulation results for each test statistic are shown in Tables 7-12 with one table for each of the six allocation targets. To simplify the presentation, the results are shown only for the four modified test statistics T_{MW} , T_{MO} , T_{MC} , T_{ML} , and the log-relative-risk test statistic T_{Risk} because they tend to have better performance than the four corresponding unmodified tests. The qualitative comparisons among test statistics, RAR methods, and allocation targets can be made based on the results in Table 6.

As shown in Table 6 (also see Tables 7-12), the worst performance can be found in the results of T_{MO} and T_{Risk} , which are often conservative with less than nominal type I error rate. T_{MW} is always slightly conservative across all simulation cases. Overall, T_{MC} is the best in attaining the correct type I error rate. T_{ML} is slightly inflated as compared with chi-square test T_{MC} . However, the simulation results not shown here indicate that T_{ML} is very robust against the unbalance in patient allocation even when sample size is 20. The comparison between different RAR methods shows that the mean type I error of GDL and SEU can usually match the correct size of tests better than other methods when T_{MC} and T_{ML} are used respectively. The type I error of DBCD is usually the largest one, except at R_{Odds} . The overall type I error of SEU is comparable with GDL.

The power comparison of different test statistics indicates that T_{Risk} is the statistic with the highest power at R_{Risk} but with a much inflated type I error. Except at R_{Risk} , T_{MC} or T_{ML} is the one with the highest power. Usually, GDL has the highest power and SEU has the lowest power among all RAR methods. DBCD and SMLE have similar power, but DBCD is more powerful in most cases. At target R_{RPW} , DL urn has the best statistical properties. On the average, the target with the lowest power achieved by test statistics is R_{Risk} . The highest overall power can usually be achieved by test statistics at R_{RSIHR} and R_{LLR} , but R_{LLR} has the disadvantage of assigning more patients to the worse treatment in some cases.

Discussion

In response-adaptive randomization, the assignment of a new patient depends on the treatment outcomes of patients previously enrolled in the trial. Delayed responses are often encountered in practice. Recently, the problem of delayed response in multi-arm generalized drop-the-loser urn and generalized Friedman's urn design is studied for both continuous and discontinuous outcomes

[11, 16, 17, 36]. It is shown that, under reasonable assumption about the delay, the asymptotic properties of adaptive design are not affected by the delay. In the present study, the primary focus is the comparison between commonly used test statistics for 2×2 tables. Based on results not shown here, a less extreme allocation with higher variation would be expected when a random delay is assumed. It is assumed that the response status of each of the patients already in the trial is available before the allocation of a new patient in our simulations evaluation.

The RAR methods simulated in the present study are aimed at assigning patients to the better treatment with probabilities higher than what otherwise would be allowed by equal randomization. The price being paid is that the sample sizes on the two comparing arms are no longer fixed, and the adaptation in patient allocation can complicate the statistical inference at the end of the trial. The properties of test statistics will change when the patient allocation ratio changes in adaptive randomization. The power of test statistics shown in the present simulation study is obtained by averaging over trials with an unknown distribution of allocation ratios. As shown in our simulation results, a large deviation from the nominal significance level of the hypothesis test can be found even under the null hypothesis. Therefore, the practice of comparing asymptotic hypothesis testing methods based solely on statistical power under the alternative hypothesis is not recommended. It is important to compare adaptive randomization methods based on both the type I error rate and the statistical power, especially when the sample size is small.

General recommendations given in the result section are based on the aggregated results across different settings. Because the performance of different test statistics, RAR methods, and allocation target are closely related to each other, recommendations under a specific scenario can

be found based on the detailed simulation results in Tables 7-12.

Based on simulation results, the Cook's correction to chi-square test statistic T_{MC} and Williams' correction to log-likelihood-ratio test T_{ML} are recommended to be used for hypothesis testing at the end of adaptive randomization. T_{MC} has good ability to attain the correct significance levels, and is relatively robust against the change of RAR method or allocation target. T_{ML} has more robust performance than T_{MC} and has higher power, but its type I error is slightly inflated as compared with T_{MC} . However, T_{ML} attains more accurate type I error than T_{MC} when the sample size is small. The original Wald-type Z test statistic T_{Wald} , which is very sensitive to patient allocation and has inflated type I error, should be avoided at small sample sizes. On the other hand, T_{MW} , the Argresti's correction to T_{Wald} , and T_{MO} the modified log-odds-ratio test are too conservative and under powered at small sample sizes.

The primary objective of current study is to compare test statistics. Since the recommended test statistics are T_{MC} and T_{ML} , the comparison between RAR methods and allocation targets are mainly based on these two selected test statistics. Among SMLE, DBCD, SEU, and GDL methods, GDL seems to be the best one due to its ability to attain the correct size of hypothesis test and comparatively higher overall power at most allocation targets. Therefore, GDL is the recommended RAR method. The sequential estimation-adjusted urn (SEU) method is comparable with GDL in controlling the type I error. However, SEU is often under powered, and the high variation in patient allocation makes it less useful in practice. The DBCD method with tuning exponent α equal to 2 is the best in targeting the true allocation ratio. When T_{MC} is the test statistic, DBCD has slightly inflated type I error and slightly lower power as compared with

GDL. Therefore, among values of α , the balances among controlling the type I error, obtaining higher power, and targeting a given allocation ratio can be reached when α is equal to 2. The simulation comparison of statistical power for different RAR methods also indicates that DL urn has the best statistical properties at R_{RPW} , mainly due to its low variation in patient allocation.

The statistical characteristics of hypothesis tests and RAR methods also depend on allocation targets. At R_{Wald} , R_{Odds} , and R_{LLR} targets, more patients could be assigned to the inferior treatment in certain parameter spaces. In contrast, R_{Risk} , R_{RPW} , and R_{RSIHR} always assign more patients to the better treatment. However, due to the more extreme allocation of R_{Risk} and R_{RPW} , both power and type I error of R_{Risk} and R_{RPW} will suffer as compared with R_{RSIHR} . On the other hand, the variation of patient allocation at R_{RSIHR} is relatively small with a stable value across all simulation scenarios. Additional, among all designs with similar power using Wald-type test statistic, R_{RSIHR} allocation ration can achieve fewer failures in the whole trial. Therefore, R_{RSIHR} is recommended among all the allocation targets in the present study.

In addition to the frequentist development on the response adaptive randomization, Bayesian decision theoretic methods has also been proposed in the context of bandit problem. The concept of “patient horizon” was brought up to include future patients to whom the current study results might be applied. The goal is to maximize the total number of success in patients enrolled in the study with or without including the patient horizon. More detailed exposition of Bayesian methods for response adaptive randomization is beyond the scope of this paper and interested readers should consult the original work on this topic [37-40].

Conclusions

The Cook's correction to chi-square test and Williams' correction to log-likelihood-ratio test are recommended for hypothesis test of RAR at small sample sizes. Among all the RAR methods compared, GDL method has better statistical properties in controlling type one error and maintaining high statistical power. The RSIHR allocation target provides a good balance between assigning more patients to the better treatment and maintaining a high overall power.

Abbreviations

RAR: Response-adaptive randomization; RPW: Randomized play-the-winner; DL: Drop-the-loser; DBCD: Doubly-adaptive biased coin design; SMLE: Sequential maximum likelihood estimation design; SEU: Sequential estimation-adjusted urn; GDL: Generalized drop-the-loser urn; RSIHR: Optimal allocation target minimizing total numbers of failure for Wald-type test statistics at fixed power; MLE: Maximum likelihood estimate.

Competing interests

The authors declare that they have no competing interests.

Authors' contribution

XMG conducted the simulation part of the study. Both XMG and JJJ participated in designing the study and writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was supported in part by grants CA16672 from the National Cancer Institute and W81XWH-06-1-0303 and W81XWH-07-1-0306 from the Department of Defense. The authors

thank Dr. Lunagomez for helpful discussions. The authors also thank Ms. Lee Ann Chastain for her help, which greatly improved the presentation of our study.

References

1. Andersen J, Faries D, Tamura R: **A randomized play-the-winner design for multi-arm clinical trials.** *Communications in Statistics-Theory and Methods* 1994, **23**:309-323.
2. Eisele JR: **The doubly adaptive biased coin design for sequential clinical trials.** *Journal of Statistical Planning and Inference* 1994, **38**:249-262.
3. Hu FF, Zhang LX: **Asymptotic properties of doubly adaptive biased coin designs for multi-treatment clinical trials.** *Annals of Statistics* 2004, **32**(1):268-301.
4. Ivanova S, Rosenberger WF, Durham S, Flournoy N: **A birth and death urn for randomized clinical trials: asymptotic methods.** *Sankhya: The Indian Journals of Statistics* 2000, **62** (B):104-118.
5. Li W, Durham SD, Flournoy N: **Randomized Pôlya urn.** In: *1996 Proceedings of the Biopharmaceutical Section of the American Statistical Association: 1997; Alexandria: American Statistical Association; 1997*: 166-170.
6. Rosenberger WF, Stallard N, Ivanova A, Harper CN, Ricks ML: **Optimal adaptive designs for binary response trials.** *Biometrics* 2001, **57**:909-913.
7. Wei LJ: **The generalized Poly's urn design for sequential medical trials.** *Annals of Statistics* 1979, **7**:291-296.
8. Wei LJ, Durham SD: **The randomized play-the-winner rule in medical trials.** *Journal of American Statistical Association* 1978, **85**:156-162.
9. Yang Y, Zhu D: **Randomized allocation with nonparametric estimation for a multi-armed bandit problem with covariates.** *Annals of Statistics* 2002, **30**:100-121.
10. Zelen M: **Play the winner rule and the controlled clinical trial.** *Journal of the American Statistical Association* 1969, **64**:131-146.
11. Zhang LX, Chan WS, Cheung SH, Hu FF: **A generalized drop-the-loser urn for clinical trials with delayed responses.** *Statistica Sinica* 2007, **17**(1):387-409.
12. Melfi VF, Page C, Geraldes M: **An adaptive randomized design with application to estimation.**

- Canadian Journal of Statistics* 2001, **29**(1):107-116.
13. Zhang LX, Hu FF, Cheung SH: **Asymptotic theorems of sequential estimation-adjusted urn models.** *Annals of Applied Probability* 2006, **16**(1):340-369.
 14. Coad DS, Ivanova A: **Bias calculations for adaptive urn designs.** *Sequential Analysis* 2001, **20**(3):91-116.
 15. Rosenberger WF, Sriram TN: **Estimation for an adaptive allocation design.** *Journal of Statistical Planning and Inference* 1997, **59**:309-319.
 16. Bai ZD, Hu FF, Rosenberger WF: **Asymptotic properties of adaptive designs for clinical trials with delayed response.** *Annals of Statistics* 2002, **30**(1):122-139.
 17. Hu FF, Zhang LJ: **Asymptotic normality of urn models for clinical trials with delayed response.** *Bernoulli* 2004, **10**:447-463.
 18. Rosenberger WF, Vidyashankar AN, Agarwal DK: **Covariate-adjusted response-adaptive designs for binary response.** *Journal of Biopharmaceutical Statistics* 2001, **11**:227-236.
 19. Hu FF, Rosenberger WF: **The Theory of Response-Adaptive Randomization in Clinical Trials.** Hoboken, New Jersey: John Wiley & Sons, Inc.; 2006.
 20. Hu FF, Rosenberger WF: **Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons.** *Journal of the American Statistical Association* 2003, **98**(463):671-678.
 21. Zhang LJ, Rosenberger WF: **Response-adaptive randomization for clinical trials with continuous outcomes.** *Biometrics* 2006, **62**(2):562-569.
 22. Hu FF, Rosenberger WF, Zhang LX: **Asymptotically best response-adaptive randomization procedures.** *Journal of Statistical Planning and Inference* 2006, **136**(6):1911-1922.
 23. Morgan CC, Coad DS: **A comparison of adaptive allocation rules for group-sequential binary response clinical trials.** *Statistics in Medicine* 2007, **26**(9):1937-1954.
 24. Guimaraes P, Palesch Y: **Power and sample size simulations for Randomized Play-the-Winner rules.** *Contemporary Clinical Trials* 2007, **28**(4):487-499.
 25. Matthews PC, Rosenberger WF: **Variance in randomized play-the-winner clinical trials.** *Statistics & Probability Letters* 1997, **35**:233-240.
 26. Bai ZD, Hu FF: **Asymptotics in randomized urn models.** *Annals of Applied Probability* 2005, **15**

- (1B):914-940.
27. Matthews PC, Rosenberger WF: **Variance in randomized play-the-winner clinical trials.** *Statistics & Probability Letters* 1997, **35**(3):233-240.
 28. Ivanova A: **A play-the-winner-type urn design with reduced variability.** *Metrika* 2003, **58**:1-13.
 29. Gart JJ: **Alternative analyses of contingency tables.** *Journal of Royal Statistical Society B* 1966, **28**:164-179.
 30. Agresti A, Caffo B: **Simple and effective confidence intervals for proportions and differences of proportions results from adding two successes and two failures.** *The American Statistician* 2000, **54**(4):280-288.
 31. Williams SS: **Improved likelihood ratio tests for complete contingency tables.** *Biometrika* 1976, **63**:33-37.
 32. Upton GJG: **A comparison of alternative tests for the 2×2 table comparative trial.** *Journal of Royal Statistical Society A* 1982, **145**:86-105.
 33. Rosenberger WF, Lachin JM: **Randomization in Clinical Trials: Theory and Practice.** New York: Wiley; 2002.
 34. Jennison C, Turnbull BW: **Group Sequential Methods with Applications to Clinical Trials.** Boca Raton: Chapman & Hall/CRC; 2000.
 35. Tymofyeyev Y, Rosenberger WF, Hu FF: **Implementing optimal allocation in sequential binary response experiments.** *Journal of American Statistical Association* 2007, **102**(477):224-234.
 36. Sun RB, Cheung SH, Zhang LX: **A generalized drop-the-loser rule for multi-treatment clinical trials.** *Journal of Statistical Planning and Inference* 2007, **137**(6):2011-2023.
 37. Berry DA, Fristedt B: **Bandit Problems.** New York: Chapman and Hall; 1985.
 38. Thompson WR: **On the likelihood that one unknown probability exceeds another in the view of the evidence of the two samples.** *Biometrika* 1933, **25**:275-294.
 39. Berry DA, Eick SG: **Adaptive assignment versus balanced randomization in clinical trials: a decision analysis.** *Statistics in Medicine* 1995, **14**:231-246.
 40. Cheng Y, Berry DA: **Optimal adaptive randomized designs for clinical trials.** *Biometrika* 2007, **94**(4):673-689.

Tables

Table 1. Summary of data from a two-arm comparative clinical trial

	Response	Failure	Margins
Treatment 1	r_1	f_1	n_1
Treatment 2	r_2	f_2	$n - n_1 = n_2$
Margins	$r_1 + r_2 = r$	$n - r = f_1 + f_2 = f$	n

n : total number of patients; n_1, n_2 : patients on treatment 1 and 2; r : total number of treatment successes; r_1, r_2 : number of successes on treatment 1 and 2.

Table 2. Test statistics

Log-Relative-Risk	$T_{Risk} = (\log(f_2 n_1 / f_1 n_2))^2 / (r_1 / n_1 f_1 + r_2 / n_2 f_2)$
Log-odds-ratio	$T_{Odds} = (\log(f_2 r_1 / f_1 r_2))^2 / (1/f_1 + 1/f_2 + 1/r_1 + 1/r_2)$
Wald-type Z	$T_{Wald} = (r_1 / n_1 - r_2 / n_2)^2 / (f_2 r_1 / n_1^3 + f_1 r_2 / n_2^3)$
Chi-Square	$T_{Chisq} = (n-1)(r_1 f_2 - r_2 f_1)^2 / r f n_1 n_2$
Log-likelihood-ratio	$T_{LLR} = 2 \cdot (r_1 \log r_1 + r_2 \log r_2 + f_1 \log f_1 + f_2 \log f_2 - r \log r - f \log f - n_1 \log n_1 - n_2 \log n_2 + n \log n)$
Gart's Correction to T_{Odds} [29]	$T_{MO} = (\log(f'_2 n'_1 / f'_1 n'_2))^2 / (r'_1 / n'_1 f'_1 + r'_2 / n'_2 f'_2)$
Agresti's Correction to T_{Wald}	$T_{MW} = (r''_1 / n''_1 - r''_2 / n''_2)^2 / (f''_2 r''_1 / n''_1{}^3 + f''_1 r''_2 / n''_2{}^3)$
Cook's Correction to T_{Chisq}	$T_{MC} = (n-1)(r_1 f_2 - r_2 f_1 - 0.5)^2 / r f n_1 n_2$
William's Correction to T_{LLR} [31]	$T_{ML} = [1 + (n_2 - r f)(n_2 - n_1 n_2) / 6 r f n_1 n_2 n]^{-1} \cdot T_{LLR}$

$$r'_1 = r_1 + 0.5, r'_2 = r_2 + 0.5, f'_1 = f_1 + 0.5, f'_2 = f_2 + 0.5, r' = r + 1, f' = f + 1, n'_1 = n_1 + 1, n'_2 = n_2 + 1, n' = n + 2$$

$$r''_1 = r_1 + 1, r''_2 = r_2 + 1, f''_1 = f_1 + 1, f''_2 = f_2 + 1, r'' = r + 2, f'' = f + 2, n''_1 = n_1 + 2, n''_2 = n_2 + 2, n'' = n + 4$$

Table 3. Allocation targets

Optimal allocation ratio (n_1/n_2) for maximizing powers	
R_{Risk}	$\sqrt{p_1 q_2 / p_2 q_1}$
$R_{Odds} /$	$\sqrt{p_2 q_2 / p_1 q_1}$
R_{Chisq}	
$R_{Wald} /$	$\sqrt{p_1 q_1 / p_2 q_2}$
R_{Neyman}	
R_{LLR}	$\{q_2 - p_2 \exp[I_1 - I_2 / (p_2 - p_1)]\} / \{-q_1 + p_1 \exp[I_1 - I_2 / (p_2 - p_1)]\}$
Other allocation targets	
$R_{RPW} /$	q_2 / q_1
R_{DL}	
R_{RSIHR}	$\sqrt{p_1 / p_2}$ (Minimize the number of failure at fixed power of T_{Wald})
$I_1 = p_1 \log(p_1) + q_1 \log(q_1), I_2 = p_2 \log(p_2) + q_2 \log(q_2)$	

Table 4. Asymptotic allocation rates on arm 1 calculated from true p_1 and p_2

p_1	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
R_{Wald} / R_{Neyman}	0.396	0.375	0.396	0.500	0.478	0.500	0.604	0.522	0.625	0.604
R_{Risk}	0.337	0.250	0.179	0.100	0.396	0.300	0.179	0.396	0.250	0.337
R_{Odds} / R_{Chisq}	0.604	0.625	0.604	0.500	0.522	0.500	0.396	0.478	0.375	0.396
R_{LLR}	0.534	0.538	0.528	0.500	0.507	0.500	0.472	0.493	0.462	0.466
R_{RSIHR}	0.366	0.309	0.274	0.250	0.436	0.396	0.366	0.458	0.427	0.469
R_{RPW} / R_{DL}	0.438	0.357	0.250	0.100	0.417	0.300	0.125	0.375	0.167	0.250

Table 5. Mean and standard deviation (in parenthesis) of allocation rate on arm 1 for $n = 30$.

Null	p_1	0.2	0.3	0.5	0.7	0.8
	p_2	0.2	0.3	0.5	0.7	0.8
Urn	RPW	0.500(0.081)	0.500(0.095)	0.500(0.129)	0.500(0.179)	0.500(0.209)
	DL	0.500(0.048)	0.500(0.058)	0.500(0.078)	0.500(0.092)	0.500(0.097)
SMLE	R_{Wald}	0.500(0.106)	0.500(0.103)	0.500(0.098)	0.500(0.103)	0.500(0.106)
	R_{Risk}	0.500(0.130)	0.500(0.134)	0.500(0.140)	0.500(0.151)	0.500(0.158)
	R_{Odds}	0.500(0.109)	0.500(0.098)	0.500(0.091)	0.500(0.099)	0.500(0.109)
	R_{LLR}	0.500(0.093)	0.500(0.092)	0.500(0.091)	0.500(0.093)	0.500(0.094)
	R_{RSIHR}	0.500(0.117)	0.500(0.116)	0.500(0.109)	0.500(0.106)	0.500(0.102)
	R_{RPW}	0.500(0.100)	0.500(0.109)	0.500(0.131)	0.500(0.166)	0.500(0.192)
DBCD	R_{Wald}	0.500(0.090)	0.500(0.075)	0.500(0.055)	0.500(0.075)	0.500(0.090)
	R_{Risk}	0.500(0.126)	0.500(0.124)	0.500(0.123)	0.500(0.127)	0.500(0.140)
	R_{Odds}	0.500(0.082)	0.500(0.061)	0.500(0.047)	0.500(0.061)	0.500(0.082)
	R_{LLR}	0.500(0.049)	0.500(0.046)	0.500(0.044)	0.500(0.047)	0.500(0.049)
	R_{RSIHR}	0.500(0.107)	0.500(0.099)	0.500(0.078)	0.500(0.060)	0.500(0.054)
	R_{RPW}	0.500(0.064)	0.500(0.074)	0.500(0.104)	0.500(0.148)	0.500(0.185)
SEU	R_{Wald}	0.500(0.113)	0.500(0.106)	0.500(0.098)	0.500(0.106)	0.500(0.114)
	R_{Risk}	0.500(0.155)	0.500(0.168)	0.500(0.195)	0.500(0.223)	0.500(0.237)
	R_{Odds}	0.500(0.101)	0.500(0.104)	0.500(0.130)	0.500(0.176)	0.500(0.196)
	R_{LLR}	0.500(0.093)	0.500(0.091)	0.500(0.091)	0.500(0.093)	0.500(0.092)
	R_{RSIHR}	0.500(0.149)	0.500(0.146)	0.500(0.131)	0.500(0.116)	0.500(0.106)
	R_{RPW}	0.500(0.135)	0.500(0.155)	0.500(0.192)	0.500(0.222)	0.500(0.233)
GDL	R_{Wald}	0.500(0.056)	0.500(0.046)	0.500(0.033)	0.500(0.047)	0.500(0.056)
	R_{Risk}	0.500(0.106)	0.500(0.114)	0.500(0.128)	0.500(0.144)	0.500(0.154)
	R_{Odds}	0.500(0.040)	0.500(0.035)	0.500(0.055)	0.500(0.090)	0.500(0.112)
	R_{LLR}	0.500(0.029)	0.500(0.026)	0.500(0.024)	0.500(0.026)	0.500(0.029)
	R_{RSIHR}	0.500(0.073)	0.500(0.070)	0.500(0.058)	0.500(0.045)	0.500(0.039)
	R_{RPW}	0.500(0.053)	0.500(0.065)	0.500(0.088)	0.500(0.116)	0.500(0.133)
Alternative	p_1	0.1	0.1	0.1	0.1	0.3
	p_2	0.3	0.5	0.7	0.9	0.5
Urn	RPW	0.444(0.080)	0.375(0.092)	0.287(0.096)	0.181(0.088)	0.430(0.109)
	DL	0.447(0.046)	0.383(0.055)	0.316(0.056)	0.249(0.053)	0.437(0.067)
SMLE	R_{Wald}	0.440(0.100)	0.424(0.098)	0.441(0.100)	0.501(0.102)	0.483(0.101)
	R_{Risk}	0.397(0.117)	0.325(0.107)	0.259(0.095)	0.186(0.079)	0.415(0.133)

	R_{Odds}	0.562(0.110)	0.577(0.107)	0.561(0.110)	0.499(0.126)	0.517(0.095)
	R_{LLR}	0.519(0.094)	0.522(0.094)	0.515(0.094)	0.499(0.095)	0.506(0.092)
	R_{RSIHR}	0.417(0.108)	0.369(0.100)	0.335(0.093)	0.312(0.087)	0.447(0.112)
	R_{RPW}	0.447(0.099)	0.384(0.105)	0.297(0.106)	0.179(0.091)	0.434(0.117)
DBCD	R_{Wald}	0.417(0.081)	0.393(0.073)	0.416(0.081)	0.499(0.095)	0.475(0.065)
	R_{Risk}	0.371(0.106)	0.285(0.086)	0.216(0.071)	0.138(0.054)	0.394(0.116)
	R_{Odds}	0.585(0.085)	0.607(0.078)	0.586(0.086)	0.499(0.110)	0.520(0.053)
	R_{LLR}	0.474(0.048)	0.468(0.046)	0.477(0.047)	0.500(0.047)	0.493(0.045)
	R_{RSIHR}	0.392(0.093)	0.332(0.077)	0.297(0.069)	0.273(0.063)	0.431(0.088)
	R_{RPW}	0.440(0.063)	0.366(0.072)	0.266(0.078)	0.129(0.064)	0.422(0.087)
SEU	R_{Wald}	0.476(0.113)	0.464(0.110)	0.473(0.113)	0.505(0.117)	0.493(0.104)
	R_{Risk}	0.433(0.143)	0.361(0.130)	0.296(0.115)	0.234(0.091)	0.440(0.166)
	R_{Odds}	0.514(0.108)	0.497(0.124)	0.462(0.143)	0.388(0.137)	0.489(0.119)
	R_{LLR}	0.510(0.093)	0.512(0.094)	0.508(0.093)	0.501(0.094)	0.503(0.092)
	R_{RSIHR}	0.461(0.143)	0.425(0.130)	0.402(0.122)	0.383(0.113)	0.475(0.136)
	R_{RPW}	0.469(0.129)	0.424(0.136)	0.367(0.135)	0.294(0.113)	0.462(0.164)
GDL	R_{Wald}	0.450(0.051)	0.437(0.046)	0.452(0.051)	0.500(0.058)	0.486(0.040)
	R_{Risk}	0.397(0.093)	0.320(0.085)	0.251(0.071)	0.181(0.055)	0.407(0.114)
	R_{Odds}	0.527(0.043)	0.508(0.053)	0.454(0.072)	0.341(0.080)	0.484(0.045)
	R_{LLR}	0.517(0.027)	0.521(0.026)	0.515(0.027)	0.500(0.028)	0.505(0.024)
	R_{RSIHR}	0.431(0.065)	0.389(0.057)	0.362(0.051)	0.342(0.047)	0.454(0.062)
	R_{RPW}	0.454(0.052)	0.399(0.063)	0.329(0.067)	0.236(0.059)	0.444(0.075)
Alternative	p_1	0.3	0.3	0.5	0.5	0.7
	p_2	0.7	0.9	0.7	0.9	0.9
Urn	RPW	0.341(0.120)	0.227(0.123)	0.411(0.147)	0.288(0.160)	0.375(0.202)
	DL	0.363(0.071)	0.290(0.066)	0.424(0.082)	0.343(0.082)	0.416(0.092)
SMLE	R_{Wald}	0.500(0.104)	0.559(0.100)	0.517(0.100)	0.576(0.099)	0.558(0.101)
	R_{Risk}	0.334(0.124)	0.238(0.109)	0.411(0.139)	0.298(0.131)	0.375(0.149)
	R_{Odds}	0.500(0.098)	0.438(0.109)	0.485(0.095)	0.423(0.107)	0.438(0.109)
	R_{LLR}	0.499(0.091)	0.483(0.093)	0.495(0.092)	0.477(0.094)	0.481(0.094)
	R_{RSIHR}	0.408(0.107)	0.378(0.103)	0.459(0.106)	0.429(0.105)	0.468(0.101)
	R_{RPW}	0.343(0.122)	0.209(0.110)	0.405(0.141)	0.255(0.136)	0.332(0.174)
DBCD	R_{Wald}	0.500(0.075)	0.585(0.081)	0.525(0.065)	0.607(0.073)	0.584(0.081)
	R_{Risk}	0.300(0.104)	0.187(0.083)	0.391(0.118)	0.250(0.108)	0.337(0.130)
	R_{Odds}	0.501(0.061)	0.413(0.086)	0.480(0.054)	0.394(0.079)	0.414(0.084)
	R_{LLR}	0.500(0.046)	0.524(0.047)	0.508(0.045)	0.532(0.046)	0.527(0.048)
	R_{RSIHR}	0.387(0.080)	0.353(0.075)	0.453(0.069)	0.417(0.066)	0.464(0.055)
	R_{RPW}	0.317(0.095)	0.157(0.082)	0.386(0.118)	0.201(0.112)	0.284(0.158)
SEU	R_{Wald}	0.502(0.106)	0.535(0.108)	0.509(0.102)	0.540(0.102)	0.532(0.108)
	R_{Risk}	0.365(0.154)	0.280(0.126)	0.437(0.197)	0.337(0.171)	0.411(0.212)
	R_{Odds}	0.453(0.134)	0.384(0.131)	0.469(0.150)	0.399(0.146)	0.438(0.177)
	R_{LLR}	0.500(0.091)	0.493(0.094)	0.498(0.093)	0.490(0.094)	0.490(0.092)
	R_{RSIHR}	0.449(0.126)	0.429(0.121)	0.479(0.124)	0.460(0.117)	0.481(0.109)
	R_{RPW}	0.408(0.162)	0.326(0.141)	0.456(0.197)	0.366(0.173)	0.423(0.208)
GDL	R_{Wald}	0.499(0.047)	0.548(0.052)	0.514(0.041)	0.562(0.046)	0.548(0.051)
	R_{Risk}	0.319(0.104)	0.220(0.078)	0.397(0.128)	0.274(0.104)	0.356(0.138)

R_{Odds}	0.431(0.064)	0.327(0.072)	0.447(0.071)	0.342(0.080)	0.390(0.102)
R_{LLR}	0.500(0.026)	0.485(0.027)	0.495(0.025)	0.479(0.026)	0.483(0.028)
R_{RSIHR}	0.423(0.056)	0.398(0.052)	0.466(0.052)	0.440(0.046)	0.472(0.038)
R_{RPW}	0.367(0.082)	0.263(0.073)	0.420(0.098)	0.303(0.092)	0.370(0.121)

Table 6. The mean and standard deviation (in parenthesis) of type I error and power.

Type I error of test statistics							
Target	Method	T_{MW}	T_{RISK}	T_{MO}	T_{MC}	T_{ML}	Row Mean
R_{Wald}	<i>SMLE</i>	4.4(1.1)	4.6(4.1)	2.0(1.4)	5.0(0.6)	6.8(0.9)	4.6(2.4)
	<i>DBCD</i>	4.3(1.4)	5.1(5.1)	1.7(1.7)	4.8(1.2)	7.2(0.8)	4.6(2.9)
	<i>SEU</i>	4.0(0.9)	3.4(2.4)	2.3(1.2)	4.8(0.2)	5.6(0.6)	4.0(1.7)
	<i>GDL</i>	4.4(0.8)	3.7(3.1)	2.1(1.6)	5.2(0.4)	6.6(1.0)	4.4(2.2)
	Mean	4.3(1.0)	4.2(3.6)	2.0(1.4)	5.0(0.7)	6.5(1.0)	4.4(2.3)
R_{Risk}	<i>SMLE</i>	4.4(1.4)	8.6(3.5)	2.4(1.8)	5.5(1.4)	6.0(1.0)	5.4(2.8)
	<i>DBCD</i>	4.6(2.0)	10.2(4.4)	2.6(2.3)	5.7(2.2)	6.5(1.4)	5.9(3.5)
	<i>SEU</i>	3.7(0.8)	7.6(2.3)	2.1(0.8)	5.4(1.3)	5.1(0.4)	4.8(2.2)
	<i>GDL</i>	4.2(1.3)	7.9(2.4)	2.4(1.9)	5.4(1.6)	5.8(1.4)	5.1(2.5)
	Mean	4.2(1.3)	8.6(3.1)	2.4(1.7)	5.5(1.5)	5.9(1.2)	5.3(2.8)
R_{Odds}	<i>SMLE</i>	3.7(0.6)	2.4(0.5)	2.9(0.5)	4.8(0.4)	4.5(0.4)	3.7(1.0)
	<i>DBCD</i>	3.6(0.7)	2.1(0.8)	3.1(0.7)	4.7(0.3)	4.1(0.2)	3.5(1.1)
	<i>SEU</i>	3.6(0.5)	3.6(0.8)	2.3(0.7)	4.7(0.3)	4.9(0.7)	3.8(1.1)
	<i>GDL</i>	3.7(0.8)	3.4(0.8)	3.0(1.1)	5.1(0.4)	4.5(0.4)	3.9(1.0)
	Mean	3.7(0.6)	2.9(0.9)	2.8(0.8)	4.9(0.4)	4.5(0.5)	3.7(1.1)
R_{LLR}	<i>SMLE</i>	4.0(0.6)	2.7(1.2)	2.7(1.0)	5.0(0.2)	5.2(0.6)	3.9(1.3)
	<i>DBCD</i>	4.2(0.8)	3.3(2.6)	2.4(1.5)	5.0(0.4)	6.1(0.8)	4.2(1.9)
	<i>SEU</i>	4.0(0.6)	2.8(1.6)	2.4(1.0)	4.9(0.2)	5.4(0.8)	3.9(1.5)
	<i>GDL</i>	3.7(0.5)	2.5(1.3)	2.7(1.2)	4.9(0.4)	5.4(0.9)	3.8(1.5)
	Mean	3.9(0.6)	2.8(1.6)	2.5(1.1)	5.0(0.3)	5.6(0.8)	4.0(1.5)
R_{RSIHR}	<i>SMLE</i>	4.2(1.1)	6.2(4.0)	2.3(1.5)	5.2(0.8)	6.1(0.7)	4.8(2.4)
	<i>DBCD</i>	4.3(1.5)	6.9(5.2)	2.0(1.6)	5.2(1.3)	6.5(1.1)	5.0(3.0)
	<i>SEU</i>	3.9(0.8)	4.8(3.4)	2.3(1.0)	4.8(0.4)	5.5(0.5)	4.3(1.9)
	<i>GDL</i>	4.3(0.9)	4.7(3.0)	2.2(1.6)	5.1(0.6)	6.1(0.9)	4.5(2.0)
	Mean	4.2(1.0)	5.7(3.8)	2.2(1.3)	5.1(0.8)	6.1(0.8)	4.6(2.3)
R_{RPW}	<i>RPW</i>	4.2(0.8)	6.2(0.5)	2.5(1.6)	5.5(1.4)	5.4(0.8)	4.8(1.7)
	<i>DL</i>	4.3(0.8)	4.8(1.0)	2.6(1.7)	5.3(0.9)	5.3(0.4)	4.5(1.4)
	<i>SMLE</i>	4.2(0.9)	6.5(0.6)	2.8(1.8)	5.4(1.6)	5.1(0.8)	4.8(1.7)
	<i>DBCD</i>	4.3(0.9)	6.7(1.0)	2.9(2.1)	5.7(1.8)	4.8(1.0)	4.9(1.9)
	<i>SEU</i>	3.8(0.6)	5.7(1.3)	2.2(0.6)	5.4(0.8)	5.1(0.6)	4.5(1.5)
	<i>GDL</i>	4.0(0.8)	5.1(0.6)	2.7(1.6)	5.2(0.7)	5.0(0.8)	4.4(1.3)
	Mean	4.1(0.8)	5.8(1.1)	2.6(1.5)	5.4(1.2)	5.1(0.7)	4.6(1.6)
Equal Allocation		4.0(0.5)	2.9(1.7)	2.4(1.0)	5.0(0.2)	5.6(0.8)	4.0(1.5)
Power of test statistics							
Target	Method	T_{MW}	T_{RISK}	T_{MO}	T_{MC}	T_{ML}	Row Mean
R_{Wald}	<i>SMLE</i>	56.6(34.1)	48.6(35.2)	48.5(36.8)	57.6(33.4)	59.4(31.9)	54.2(33.2)
	<i>DBCD</i>	56.9(34.4)	49.5(35.9)	48.0(37.6)	57.7(33.9)	60.2(31.8)	54.5(33.7)
	<i>SEU</i>	56.0(34.0)	47.7(34.8)	49.6(36.1)	57.5(33.0)	58.4(32.3)	53.8(32.9)
	<i>GDL</i>	57.3(34.0)	50.0(36.2)	50.6(36.9)	58.4(33.2)	60.0(32.0)	55.3(33.3)
	Mean	56.7(32.8)	49.0(34.2)	49.2(35.4)	57.8(32.1)	59.5(30.7)	54.4(33.0)
R_{Risk}	<i>SMLE</i>	53.4(33.2)	57.9(31.5)	45.4(35.2)	56.2(32.7)	55.1(31.1)	53.6(31.7)
	<i>DBCD</i>	53.3(33.4)	60.0(30.5)	43.7(36.0)	56.5(32.9)	55.0(31.1)	53.7(31.9)
	<i>SEU</i>	52.5(32.8)	55.3(32.2)	45.9(34.1)	55.2(32.1)	54.2(31.2)	52.6(31.3)
	<i>GDL</i>	53.2(33.3)	58.1(31.6)	45.8(35.8)	56.5(32.6)	55.2(31.7)	53.8(31.9)
	Mean	53.1(31.9)	57.8(30.3)	45.2(33.9)	56.1(31.3)	54.9(30.1)	53.4(31.5)
R_{Odds}	<i>SMLE</i>	54.6(33.9)	47.1(34.3)	52.1(34.9)	57.6(32.6)	56.4(32.9)	53.6(32.5)
	<i>DBCD</i>	54.8(34.2)	47.3(35.2)	53.4(34.5)	57.8(32.7)	56.5(33.4)	53.9(32.8)
	<i>SEU</i>	54.8(33.5)	50.8(33.8)	50.4(34.8)	57.5(32.5)	56.6(32.2)	54.0(32.1)
	<i>GDL</i>	54.6(34.2)	53.0(34.6)	52.5(35.0)	58.1(32.7)	56.8(33.0)	55.0(32.5)
	Mean	54.7(32.6)	49.5(33.2)	52.1(33.4)	57.8(31.4)	56.6(31.6)	54.1(32.3)

R_{LLR}	<i>SMLE</i>	55.9(33.9)	48.4(35.0)	51.6(35.6)	58.0(32.8)	58.0(32.6)	54.4(32.8)
	<i>DBCD</i>	57.2(34.0)	49.9(35.9)	51.4(36.6)	58.6(33.1)	60.0(32.2)	55.4(33.2)
	<i>SEU</i>	56.1(33.9)	48.5(34.8)	51.2(35.7)	58.1(32.8)	58.2(32.5)	54.4(32.8)
	<i>GDL</i>	56.4(34.1)	50.4(35.8)	53.1(35.9)	58.9(33.1)	59.5(32.5)	55.7(33.1)
	Mean	56.4(32.6)	49.3(34.0)	51.8(34.6)	58.4(31.7)	58.9(31.2)	55.0(32.7)
R_{RSIHR}	<i>SMLE</i>	56.0(33.9)	54.8(33.7)	48.7(36.4)	57.5(33.2)	58.4(32.0)	55.1(32.6)
	<i>DBCD</i>	56.8(34.0)	56.3(33.4)	48.2(37.0)	58.2(33.2)	59.4(31.8)	55.7(32.8)
	<i>SEU</i>	54.5(33.8)	50.5(34.5)	48.6(35.8)	56.4(33.0)	56.6(32.4)	53.3(32.7)
	<i>GDL</i>	57.4(33.7)	54.4(34.5)	50.6(36.6)	58.7(33.0)	59.7(32.1)	56.2(32.8)
	Mean	56.2(32.6)	54.0(32.8)	49.0(35.0)	57.7(31.8)	58.5(30.8)	55.1(32.5)
R_{RPW}	<i>RPW</i>	52.4(32.3)	55.9(32.1)	46.3(34.1)	55.8(32.1)	52.9(30.1)	52.7(31.0)
	<i>DL</i>	56.0(33.5)	55.9(33.4)	50.0(36.1)	58.2(32.6)	57.4(32.5)	55.5(32.4)
	<i>SMLE</i>	51.7(32.3)	56.2(31.8)	46.7(33.7)	55.7(31.9)	51.7(30.2)	52.4(30.9)
	<i>DBCD</i>	51.2(31.8)	57.3(31.2)	47.0(34.1)	56.0(31.5)	48.3(29.2)	52.0(30.6)
	<i>SEU</i>	54.0(33.1)	54.0(32.7)	48.3(34.4)	56.7(32.1)	55.9(31.7)	53.8(31.6)
	<i>GDL</i>	54.6(33.5)	56.0(33.0)	50.2(35.3)	57.8(32.4)	56.4(32.3)	55.0(32.0)
	Mean	53.3(31.4)	55.9(31.0)	48.1(33.2)	56.7(30.7)	53.8(29.8)	53.5(31.2)
	Equal Allocation	56.2(33.9)	48.5(35.0)	50.9(35.9)	58.1(32.9)	58.4(32.4)	54.4(32.9)

Mean values are calculated by averaging simulation results over the five null cases and the ten alternative cases of simulation scenarios listed in Tables 7-12. All results have been multiplied by 100% ($\alpha = 0.05$, $n = 30$).

Table 7. Power and type I error at R_{Wald} (alpha = 0.05, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
SMLE	T_{MW}	0.031	0.048	0.056	0.050	0.033	0.196	0.674	0.953	0.999	0.201	0.600	0.950	0.203	0.680	0.202
	T_{Risk}	0.102	0.072	0.039	0.014	0.003	0.326	0.693	0.940	0.996	0.181	0.501	0.798	0.113	0.288	0.024
	T_{MO}	0.007	0.022	0.041	0.024	0.007	0.063	0.492	0.928	0.999	0.162	0.563	0.923	0.161	0.495	0.069
	T_{MC}	0.044	0.052	0.056	0.055	0.044	0.231	0.689	0.954	0.999	0.203	0.601	0.952	0.205	0.693	0.235
	T_{ML}	0.074	0.066	0.055	0.067	0.079	0.308	0.709	0.954	0.999	0.203	0.595	0.951	0.205	0.711	0.309
DBCD	T_{MW}	0.029	0.050	0.057	0.052	0.026	0.186	0.685	0.957	0.999	0.212	0.607	0.958	0.206	0.696	0.191
	T_{Risk}	0.120	0.085	0.041	0.008	0.001	0.361	0.721	0.954	0.998	0.204	0.524	0.811	0.109	0.257	0.010
	T_{MO}	0.004	0.017	0.045	0.017	0.003	0.041	0.462	0.933	0.999	0.169	0.587	0.934	0.164	0.475	0.042
	T_{MC}	0.037	0.056	0.058	0.056	0.034	0.211	0.696	0.958	0.999	0.215	0.607	0.959	0.208	0.706	0.215
	T_{ML}	0.077	0.074	0.059	0.073	0.077	0.311	0.718	0.958	0.999	0.217	0.607	0.959	0.210	0.727	0.315
SEU	T_{MW}	0.031	0.045	0.048	0.044	0.030	0.200	0.655	0.946	0.999	0.190	0.583	0.948	0.191	0.675	0.213
	T_{Risk}	0.067	0.048	0.033	0.016	0.006	0.259	0.646	0.922	0.991	0.154	0.486	0.812	0.114	0.342	0.046
	T_{MO}	0.013	0.026	0.039	0.027	0.011	0.094	0.522	0.921	0.999	0.158	0.553	0.926	0.157	0.533	0.095
	T_{MC}	0.046	0.051	0.049	0.050	0.046	0.248	0.675	0.949	0.999	0.195	0.585	0.950	0.195	0.698	0.258
	T_{ML}	0.062	0.055	0.047	0.055	0.062	0.285	0.683	0.947	0.999	0.190	0.577	0.949	0.193	0.710	0.305
GDL	T_{MW}	0.036	0.051	0.051	0.049	0.034	0.223	0.696	0.954	1.000	0.195	0.601	0.958	0.200	0.692	0.214
	T_{Risk}	0.075	0.060	0.040	0.010	0.001	0.309	0.703	0.949	0.999	0.184	0.543	0.868	0.124	0.304	0.015
	T_{MO}	0.007	0.022	0.046	0.023	0.006	0.077	0.549	0.937	0.999	0.167	0.588	0.945	0.169	0.547	0.077
	T_{MC}	0.048	0.057	0.051	0.055	0.047	0.260	0.708	0.955	1.000	0.198	0.602	0.960	0.204	0.705	0.253
	T_{ML}	0.074	0.064	0.052	0.063	0.076	0.319	0.721	0.956	1.000	0.200	0.602	0.960	0.205	0.720	0.314

For each RAR methods, the results of the following 5 test statistics are shown: Agresti's correction to Wald-type Z test T_{MW} , log-relative-risk test T_{Risk} , Gart's correction to log-odds-ratio test T_{MO} , Cook's correction to chi-square test T_{MC} , and Williams' correction log-likelihood-ratio test T_{ML} .

Table 8. Power and type I error at R_{Risk} (alpha = 0.05, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
SMLE	T_{MW}	0.024	0.045	0.061	0.051	0.041	0.156	0.615	0.923	0.990	0.185	0.560	0.898	0.189	0.611	0.214
	T_{Risk}	0.136	0.105	0.078	0.061	0.050	0.363	0.716	0.945	0.997	0.230	0.588	0.923	0.206	0.612	0.210
	T_{MO}	0.002	0.008	0.032	0.039	0.040	0.022	0.278	0.792	0.988	0.096	0.466	0.903	0.157	0.615	0.220
	T_{MC}	0.033	0.047	0.060	0.064	0.068	0.177	0.615	0.923	0.996	0.183	0.570	0.939	0.202	0.701	0.316
	T_{ML}	0.069	0.071	0.061	0.049	0.051	0.278	0.659	0.921	0.975	0.195	0.543	0.883	0.179	0.621	0.253
DBCD	T_{MW}	0.018	0.046	0.072	0.054	0.042	0.134	0.617	0.931	0.993	0.198	0.565	0.896	0.199	0.586	0.207
	T_{Risk}	0.166	0.123	0.091	0.066	0.062	0.402	0.744	0.951	0.998	0.253	0.606	0.926	0.225	0.649	0.243
	T_{MO}	0.001	0.003	0.030	0.046	0.049	0.004	0.164	0.746	0.994	0.074	0.457	0.904	0.158	0.623	0.248
	T_{MC}	0.023	0.047	0.070	0.068	0.077	0.148	0.612	0.928	0.998	0.193	0.575	0.940	0.218	0.707	0.327
	T_{ML}	0.071	0.083	0.071	0.050	0.050	0.278	0.665	0.928	0.979	0.207	0.549	0.880	0.184	0.596	0.240
SEU	T_{MW}	0.026	0.039	0.045	0.043	0.032	0.172	0.598	0.903	0.988	0.178	0.537	0.888	0.183	0.606	0.198
	T_{Risk}	0.105	0.092	0.075	0.059	0.049	0.307	0.686	0.935	0.996	0.201	0.546	0.903	0.186	0.581	0.193
	T_{MO}	0.009	0.018	0.029	0.027	0.023	0.062	0.372	0.794	0.986	0.121	0.468	0.887	0.146	0.582	0.176
	T_{MC}	0.041	0.044	0.050	0.064	0.070	0.209	0.605	0.903	0.994	0.178	0.542	0.922	0.194	0.681	0.289
	T_{ML}	0.057	0.052	0.047	0.049	0.048	0.266	0.640	0.900	0.981	0.183	0.526	0.879	0.178	0.624	0.245
GDL	T_{MW}	0.023	0.043	0.059	0.047	0.038	0.168	0.617	0.929	0.993	0.182	0.558	0.902	0.196	0.580	0.195
	T_{Risk}	0.113	0.092	0.076	0.062	0.053	0.347	0.720	0.950	0.998	0.227	0.593	0.928	0.220	0.617	0.213
	T_{MO}	0.001	0.006	0.031	0.040	0.042	0.016	0.283	0.831	0.994	0.094	0.473	0.908	0.161	0.604	0.220
	T_{MC}	0.030	0.047	0.058	0.064	0.070	0.194	0.618	0.928	0.998	0.180	0.567	0.943	0.214	0.696	0.311
	T_{ML}	0.077	0.068	0.058	0.044	0.045	0.292	0.653	0.927	0.990	0.189	0.540	0.901	0.182	0.606	0.236

Table 9. Power and type I error at R_{Odds} (alpha = 0.05, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
SMLE	T_{MW}	0.030	0.040	0.042	0.040	0.031	0.202	0.630	0.935	0.998	0.178	0.562	0.939	0.174	0.637	0.205
	T_{Risk}	0.022	0.023	0.030	0.026	0.017	0.143	0.502	0.857	0.984	0.128	0.475	0.884	0.129	0.497	0.112
	T_{MO}	0.024	0.031	0.036	0.031	0.023	0.163	0.587	0.926	0.999	0.154	0.536	0.929	0.151	0.598	0.167
	T_{MC}	0.053	0.048	0.043	0.047	0.052	0.283	0.682	0.946	0.999	0.184	0.566	0.947	0.180	0.690	0.285
	T_{ML}	0.048	0.045	0.040	0.044	0.049	0.266	0.662	0.938	0.998	0.174	0.551	0.941	0.171	0.672	0.270
DBCD	T_{MW}	0.029	0.040	0.044	0.040	0.028	0.191	0.632	0.940	0.999	0.180	0.572	0.941	0.178	0.644	0.198
	T_{Risk}	0.011	0.018	0.032	0.026	0.018	0.085	0.448	0.864	0.994	0.120	0.490	0.906	0.141	0.547	0.134
	T_{MO}	0.026	0.033	0.042	0.031	0.024	0.178	0.609	0.934	0.999	0.165	0.555	0.933	0.161	0.619	0.185
	T_{MC}	0.052	0.046	0.045	0.046	0.048	0.280	0.688	0.948	0.999	0.185	0.573	0.949	0.181	0.696	0.284
	T_{ML}	0.040	0.043	0.043	0.043	0.038	0.244	0.667	0.945	0.999	0.178	0.565	0.944	0.174	0.680	0.252
SEU	T_{MW}	0.032	0.041	0.043	0.037	0.030	0.207	0.647	0.935	0.996	0.183	0.562	0.924	0.186	0.636	0.204
	T_{Risk}	0.047	0.040	0.035	0.032	0.028	0.214	0.605	0.903	0.993	0.152	0.503	0.894	0.140	0.528	0.146
	T_{MO}	0.014	0.026	0.032	0.023	0.020	0.127	0.540	0.900	0.995	0.148	0.520	0.914	0.150	0.587	0.159
	T_{MC}	0.049	0.047	0.043	0.047	0.052	0.268	0.676	0.938	0.998	0.187	0.564	0.945	0.191	0.695	0.284
	T_{ML}	0.059	0.049	0.042	0.044	0.049	0.285	0.677	0.935	0.995	0.182	0.551	0.922	0.183	0.665	0.268
GDL	T_{MW}	0.029	0.037	0.049	0.041	0.030	0.203	0.657	0.943	0.999	0.167	0.573	0.929	0.178	0.617	0.192
	T_{Risk}	0.024	0.032	0.046	0.035	0.031	0.183	0.625	0.936	0.999	0.158	0.560	0.922	0.165	0.583	0.166
	T_{MO}	0.013	0.026	0.043	0.034	0.033	0.124	0.587	0.930	0.999	0.150	0.552	0.928	0.161	0.619	0.204
	T_{MC}	0.051	0.047	0.050	0.050	0.058	0.281	0.700	0.948	0.999	0.177	0.579	0.949	0.187	0.695	0.298
	T_{ML}	0.050	0.047	0.046	0.039	0.043	0.282	0.700	0.947	0.999	0.176	0.563	0.933	0.169	0.652	0.258

Table 10. Power and type I error at R_{LLR} (alpha = 0.05, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
SMLE	T_{MW}	0.034	0.043	0.046	0.044	0.031	0.212	0.659	0.946	0.999	0.187	0.575	0.948	0.182	0.667	0.218
	T_{Risk}	0.039	0.034	0.033	0.022	0.008	0.203	0.597	0.911	0.995	0.146	0.490	0.869	0.124	0.432	0.072
	T_{MO}	0.018	0.029	0.040	0.031	0.017	0.129	0.577	0.931	0.999	0.162	0.549	0.934	0.156	0.587	0.133
	T_{MC}	0.052	0.050	0.046	0.052	0.051	0.274	0.692	0.951	0.999	0.192	0.578	0.953	0.185	0.700	0.278
	T_{ML}	0.060	0.050	0.044	0.051	0.057	0.289	0.691	0.948	0.999	0.186	0.567	0.950	0.181	0.698	0.289
DBCD	T_{MW}	0.036	0.047	0.050	0.045	0.031	0.223	0.688	0.957	0.999	0.192	0.591	0.956	0.192	0.697	0.225
	T_{Risk}	0.063	0.049	0.037	0.012	0.001	0.278	0.686	0.947	0.998	0.171	0.528	0.872	0.129	0.356	0.026
	T_{MO}	0.010	0.028	0.046	0.026	0.009	0.094	0.569	0.946	0.999	0.169	0.579	0.942	0.171	0.580	0.094
	T_{MC}	0.050	0.055	0.051	0.052	0.044	0.265	0.710	0.959	0.999	0.197	0.592	0.959	0.197	0.715	0.267
	T_{ML}	0.071	0.062	0.051	0.057	0.066	0.315	0.727	0.960	0.999	0.198	0.591	0.959	0.199	0.733	0.316
SEU	T_{MW}	0.034	0.043	0.046	0.043	0.033	0.215	0.665	0.947	0.999	0.187	0.581	0.947	0.186	0.671	0.214
	T_{Risk}	0.047	0.038	0.031	0.018	0.007	0.226	0.617	0.915	0.995	0.148	0.492	0.854	0.125	0.414	0.063
	T_{MO}	0.016	0.027	0.038	0.028	0.013	0.124	0.573	0.931	0.999	0.161	0.553	0.929	0.157	0.574	0.123
	T_{MC}	0.052	0.049	0.047	0.050	0.050	0.276	0.696	0.952	0.999	0.191	0.583	0.951	0.191	0.701	0.270
	T_{ML}	0.063	0.051	0.044	0.052	0.061	0.294	0.696	0.949	0.999	0.186	0.573	0.948	0.186	0.701	0.292
GDL	T_{MW}	0.033	0.037	0.043	0.038	0.032	0.230	0.670	0.950	1.000	0.178	0.585	0.956	0.177	0.675	0.215
	T_{Risk}	0.035	0.032	0.036	0.018	0.005	0.230	0.645	0.937	0.999	0.151	0.537	0.905	0.139	0.449	0.049
	T_{MO}	0.016	0.030	0.043	0.031	0.014	0.139	0.614	0.945	1.000	0.172	0.582	0.951	0.172	0.612	0.127
	T_{MC}	0.052	0.050	0.044	0.048	0.053	0.293	0.719	0.955	1.000	0.189	0.588	0.960	0.186	0.722	0.275
	T_{ML}	0.063	0.051	0.044	0.049	0.064	0.322	0.722	0.955	1.000	0.189	0.587	0.960	0.187	0.728	0.302

Table 11. Power and type I error at R_{RSIHR} (alpha = 0.05, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
SMLE	T_{MW}	0.028	0.045	0.056	0.048	0.035	0.174	0.648	0.944	0.999	0.192	0.588	0.946	0.202	0.678	0.228
	T_{Risk}	0.118	0.085	0.058	0.034	0.018	0.343	0.712	0.950	0.999	0.207	0.568	0.910	0.172	0.515	0.102
	T_{MO}	0.004	0.012	0.040	0.034	0.023	0.037	0.397	0.890	0.998	0.130	0.538	0.936	0.170	0.616	0.156
	T_{MC}	0.038	0.049	0.056	0.057	0.057	0.200	0.657	0.945	0.999	0.192	0.591	0.953	0.208	0.718	0.290
	T_{ML}	0.070	0.065	0.056	0.054	0.062	0.291	0.685	0.945	0.998	0.196	0.579	0.946	0.197	0.705	0.301
DBCD	T_{MW}	0.020	0.050	0.057	0.050	0.038	0.157	0.654	0.948	0.999	0.201	0.605	0.956	0.217	0.700	0.242
	T_{Risk}	0.138	0.103	0.062	0.030	0.013	0.383	0.732	0.953	0.999	0.227	0.594	0.922	0.186	0.534	0.097
	T_{MO}	0.001	0.007	0.038	0.034	0.020	0.017	0.323	0.887	0.999	0.123	0.554	0.942	0.185	0.628	0.159
	T_{MC}	0.028	0.056	0.057	0.057	0.060	0.183	0.662	0.948	0.999	0.202	0.607	0.959	0.221	0.733	0.304
	T_{ML}	0.074	0.079	0.057	0.052	0.064	0.293	0.693	0.948	0.999	0.208	0.593	0.954	0.207	0.726	0.317
SEU	T_{MW}	0.029	0.039	0.050	0.044	0.033	0.181	0.626	0.930	0.998	0.178	0.559	0.932	0.182	0.653	0.214
	T_{Risk}	0.095	0.070	0.044	0.024	0.010	0.275	0.650	0.926	0.996	0.163	0.512	0.875	0.137	0.449	0.071
	T_{MO}	0.014	0.021	0.037	0.028	0.016	0.075	0.466	0.892	0.997	0.137	0.521	0.921	0.152	0.574	0.128
	T_{MC}	0.044	0.045	0.050	0.053	0.049	0.225	0.642	0.932	0.998	0.181	0.562	0.945	0.189	0.696	0.271
	T_{ML}	0.058	0.053	0.050	0.052	0.062	0.268	0.657	0.929	0.997	0.178	0.548	0.934	0.179	0.684	0.289
GDL	T_{MW}	0.031	0.048	0.052	0.050	0.036	0.206	0.682	0.951	1.000	0.197	0.610	0.961	0.212	0.690	0.235
	T_{Risk}	0.084	0.065	0.050	0.026	0.009	0.321	0.715	0.952	1.000	0.201	0.591	0.919	0.173	0.495	0.076
	T_{MO}	0.002	0.016	0.042	0.034	0.017	0.047	0.476	0.923	1.000	0.147	0.577	0.947	0.186	0.613	0.142
	T_{MC}	0.040	0.052	0.052	0.056	0.053	0.228	0.689	0.952	1.000	0.198	0.611	0.964	0.216	0.721	0.289
	T_{ML}	0.074	0.062	0.051	0.055	0.063	0.301	0.707	0.952	1.000	0.199	0.602	0.962	0.207	0.722	0.316

Table 12. Power and type I error at R_{RPW} ($\alpha = 0.05$, $n = 30$).

p_1		0.200	0.300	0.500	0.700	0.800	0.100	0.100	0.100	0.100	0.300	0.300	0.300	0.500	0.500	0.700
p_2		0.200	0.300	0.500	0.700	0.800	0.300	0.500	0.700	0.900	0.500	0.700	0.900	0.700	0.900	0.900
RPW	T_{MW}	0.031	0.039	0.050	0.050	0.042	0.191	0.631	0.918	0.966	0.166	0.538	0.859	0.183	0.585	0.204
	T_{Risk}	0.071	0.058	0.059	0.061	0.060	0.287	0.683	0.939	0.993	0.193	0.565	0.905	0.197	0.607	0.216
	T_{MO}	0.004	0.012	0.032	0.038	0.039	0.047	0.410	0.840	0.967	0.105	0.467	0.867	0.151	0.584	0.196
	T_{MC}	0.045	0.042	0.050	0.063	0.075	0.227	0.640	0.921	0.988	0.167	0.546	0.914	0.196	0.680	0.301
	T_{ML}	0.067	0.050	0.049	0.049	0.053	0.288	0.661	0.916	0.931	0.172	0.523	0.820	0.173	0.573	0.235
DL	T_{MW}	0.032	0.043	0.052	0.050	0.040	0.208	0.658	0.944	0.998	0.183	0.586	0.939	0.204	0.658	0.219
	T_{Risk}	0.057	0.051	0.055	0.048	0.032	0.273	0.679	0.947	0.998	0.192	0.588	0.935	0.199	0.612	0.164
	T_{MO}	0.003	0.013	0.038	0.041	0.033	0.047	0.464	0.906	0.998	0.123	0.527	0.934	0.172	0.641	0.193
	T_{MC}	0.043	0.045	0.052	0.062	0.064	0.237	0.662	0.944	0.999	0.184	0.592	0.956	0.216	0.723	0.307
	T_{ML}	0.058	0.050	0.050	0.049	0.056	0.275	0.672	0.943	0.998	0.183	0.567	0.940	0.188	0.688	0.283
SMLE	T_{MW}	0.027	0.040	0.048	0.049	0.044	0.188	0.626	0.921	0.968	0.167	0.537	0.848	0.175	0.550	0.195
	T_{Risk}	0.073	0.062	0.058	0.063	0.072	0.283	0.678	0.936	0.993	0.193	0.563	0.910	0.196	0.617	0.247
	T_{MO}	0.006	0.012	0.031	0.040	0.049	0.054	0.409	0.840	0.969	0.108	0.463	0.864	0.148	0.584	0.229
	T_{MC}	0.039	0.044	0.049	0.061	0.079	0.226	0.636	0.922	0.989	0.168	0.547	0.911	0.190	0.671	0.315
	T_{ML}	0.064	0.054	0.046	0.046	0.047	0.287	0.659	0.917	0.925	0.171	0.519	0.794	0.165	0.528	0.200
DBCD	T_{MW}	0.031	0.037	0.053	0.049	0.044	0.202	0.635	0.929	0.969	0.181	0.529	0.813	0.173	0.503	0.192
	T_{Risk}	0.063	0.054	0.065	0.072	0.081	0.290	0.685	0.942	0.994	0.202	0.572	0.911	0.209	0.640	0.285
	T_{MO}	0.003	0.010	0.033	0.043	0.054	0.041	0.407	0.866	0.981	0.110	0.460	0.856	0.146	0.573	0.257
	T_{MC}	0.041	0.040	0.054	0.067	0.083	0.236	0.640	0.930	0.990	0.181	0.543	0.905	0.195	0.660	0.325
	T_{ML}	0.061	0.048	0.052	0.042	0.036	0.289	0.661	0.925	0.857	0.183	0.511	0.696	0.160	0.407	0.144
SEU	T_{MW}	0.033	0.040	0.047	0.041	0.032	0.204	0.633	0.924	0.994	0.183	0.553	0.908	0.185	0.618	0.199
	T_{Risk}	0.076	0.059	0.058	0.048	0.043	0.278	0.664	0.929	0.996	0.183	0.529	0.899	0.170	0.564	0.182
	T_{MO}	0.012	0.021	0.028	0.027	0.024	0.100	0.467	0.855	0.993	0.130	0.493	0.900	0.143	0.578	0.169
	T_{MC}	0.051	0.047	0.050	0.059	0.065	0.251	0.652	0.925	0.997	0.186	0.556	0.933	0.197	0.686	0.286
	T_{ML}	0.062	0.051	0.048	0.047	0.049	0.293	0.671	0.923	0.992	0.185	0.541	0.904	0.183	0.642	0.251
GDL	T_{MW}	0.032	0.045	0.049	0.045	0.032	0.216	0.658	0.937	0.998	0.171	0.576	0.916	0.192	0.602	0.196
	T_{Risk}	0.056	0.053	0.053	0.050	0.042	0.281	0.681	0.942	0.998	0.180	0.586	0.927	0.196	0.615	0.197
	T_{MO}	0.004	0.017	0.036	0.040	0.037	0.066	0.480	0.900	0.998	0.122	0.525	0.918	0.165	0.622	0.219
	T_{MC}	0.044	0.049	0.050	0.058	0.061	0.250	0.666	0.939	0.999	0.173	0.584	0.948	0.206	0.700	0.314
	T_{ML}	0.061	0.054	0.047	0.044	0.043	0.294	0.681	0.937	0.998	0.175	0.560	0.920	0.179	0.639	0.256

Comparison and Optimization of Multiplexed Quantum Dot-Based Immunohistofluorescence

Donghai Huang^{1,2}, Xianghong Peng¹, Ling Su¹, Dongsheng Wang¹, Fadlo R. Khuri¹, Dong M. Shin¹, and Zhuo (Georgia) Chen¹ (✉)

¹ Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia 30322, USA

² Department of Otolaryngology-Head and Neck Surgery, Xiang-Ya Hospital, Central South University, Changsha, Hunan 410008, China

Received: 25 October 2009 / Revised: 25 November 2009 / Accepted: 25 November 2009

© The Author(s) 2010. This article is published with open access at Springerlink.com

ABSTRACT

Nanoparticle quantum dots (QDs) are ideal materials for multiplexed biomarker detection, localization, and quantification. Both direct and indirect methods are available for QD-based immunohistofluorescence (QD-IHF) staining; the direct method, however, has been considered laborious and costly. In this study, we optimized and compared the indirect QD-IHF single staining procedure using QD-secondary antibody conjugates and QD-streptavidin conjugates. Problems associated with sequential multiplex staining were identified quantitatively. A method using a QD cocktail solution was developed allowing simultaneous staining with three antibodies against E-cadherin, epidermal growth factor receptor and β -catenin in formalin-fixed and paraffin-embedded (FFPE) tissues. The expression of each biomarker was quantified by using the cocktail and the sequential methods. Comparison of the two methods demonstrated that the cocktail method provided more consistent and stable QD signals for each multiplexed biomarker than the sequential method, and provides a convenient tool for multiplexing biomarkers in both research and clinical applications.

KEYWORDS

Quantum dots, immunohistofluorescence, nanocrystals, spectral imaging

Introduction

In recent years, nanotechnology has developed rapidly and been used in molecular detection, imaging, diagnostics, and therapeutics in the cancer field [1, 2]. Quantum dots (QDs) are nanoscale particles made from inorganic semiconductors that can produce different fluorescence signals depending on their size and components. Compared with organic dyes, QDs have superior signal brightness and photostability, relatively long excited-state lifetime, and optimized

signal-to-background ratios [3]. QDs can be covalently linked to biological molecules such as peptides, proteins, and nucleic acids, as well as streptavidin [4, 5]. Due to their long excitation time and narrow emission spectra, QDs can be excited simultaneously through one appropriate excitation source. Together these properties render QDs ideal for multiplexed biological imaging and they have been used for both molecular and cellular labeling [3–7].

Many researchers reported that QDs can immunostain more than three biomarkers in formalin-fixed

Address correspondence to gzchen@emory.edu



Springer

paraffin-embedded (FFPE) tissues using QD-based immunohistofluorescence (QD-IHF) [8–11]. To date, several different staining procedures have been utilized, including direct and indirect staining, such as QDs linked to primary antibodies and QDs linked to secondary antibodies or streptavidin, respectively [9, 10, 12, 13]. Although the direct staining method (QDs linked directly to a primary antibody) is straightforward, the affinity of some primary antibodies may be reduced during the QD conjugation process. The conformation and function of the primary antibody may be changed and its binding properties are likely altered by covalent modifications at either $-\text{NH}_2$ or $-\text{COOH}$ sites [9, 14]. Furthermore, the reagent costs are considerable because each conjugation reaction requires up to 300 μg of antibody (Invitrogen protocol) and the yield of QD-antibody conjugates is usually low. Since each primary antibody is covalently conjugated to just one type of QD, changing the antibody for a particular QD probe is not possible once the conjugation is completed. Many researchers have abandoned the direct staining method since these problems can be avoided by indirect QD staining methods.

The main advantages of indirect QD staining are its flexibility, lower costs, and the reduced constraints on primary antibodies. Although many studies have described detailed protocols for tissue specimen preparation, multicolor QD staining, and image processing [8, 9, 15], these have not provided detailed discussion or quantitative analysis of the optimization of their multiplexed biomarker staining procedures. In this study, we compared multiple QD staining in a sequential order with that in a simultaneous combination while using different methods—QD-secondary antibody conjugates and QD-streptavidin conjugates. QD staining of three biomarkers, epidermal growth factor receptor (EGFR), E-cadherin (E-cad), and β -catenin (β -cat), was examined and quantitatively evaluated with these staining methods for each of the tested biomarkers.

1. Experimental

Using an institutional review board-approved consent for tissue acquisition, tissue samples for this study

were obtained from surgical specimens from patients who were diagnosed at Emory University Hospital with squamous cell carcinoma of the head and neck (SCCHN). After a routine process to generate FFPE samples, the blocks were sectioned to 4 μm each and mounted on coated slides. Each sample was analyzed by hematoxylin and eosin (H&E) staining. Before QD-IHF staining, we confirmed that the primary antibodies were suitable for immunohistochemistry (IHC) and also selected FFPE samples which were strongly positive for staining of the primary antibodies as positive control slides.

We selected mouse anti-human E-cad, rabbit anti-human EGFR, and goat anti-human β -cat as the primary antibodies, since these antigens are strongly expressed in SCCHN tissues. We found previously that expression and localization of E-cad and EGFR correlated with metastasis and poor prognosis [16]. QD-secondary antibody conjugates (QD-2nd Ab) and QD-streptavidin conjugates (QD-streptavidin) were selected as follows: Qdots® 565 goat F(ab')₂ anti-mouse immunoglobulin G (IgG) conjugate, Qdots® 605 goat F(ab')₂ anti-rabbit IgG conjugate, Qdots® 655 rabbit F(ab')₂ anti-goat IgG conjugate, and Qdots® streptavidin conjugate (565, 605, 655).

The QD-IHF procedure with single QD-2nd Ab was carried out as follows (shown in the cartoon in Fig. 1(a)). After deparaffinization and rehydration, antigen retrieval was performed by heating with citric acid (10 mmol/L, pH 6.0) in a microwave to 95 °C for 10 min. The tissue slides were blocked with 5% normal goat serum (Dako) for 10 min before the primary antibody incubation (E-cad 1:2000 dilution, EGFR 1:150 dilution, or β -cat 1:2000 dilution) for 1 h at 37 °C. Following three washes with phosphate-buffered saline (PBS) (5 min each), the slides were incubated with QD [QD565 goat F(ab')₂ anti-mouse IgG conjugate, QD605 goat F(ab')₂ anti-rabbit IgG conjugate, or QD655 rabbit F(ab')₂ anti-goat IgG conjugate, accordingly] in 6% bovine serum albumin (BSA) for 1 h at 37 °C. After washing three times with PBS, the nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI) (Invitrogen, Carlsbad, CA, USA). The slides were mounted with Cytoseal™ 60 mounting medium (Richard-Allan Scientific, MI).

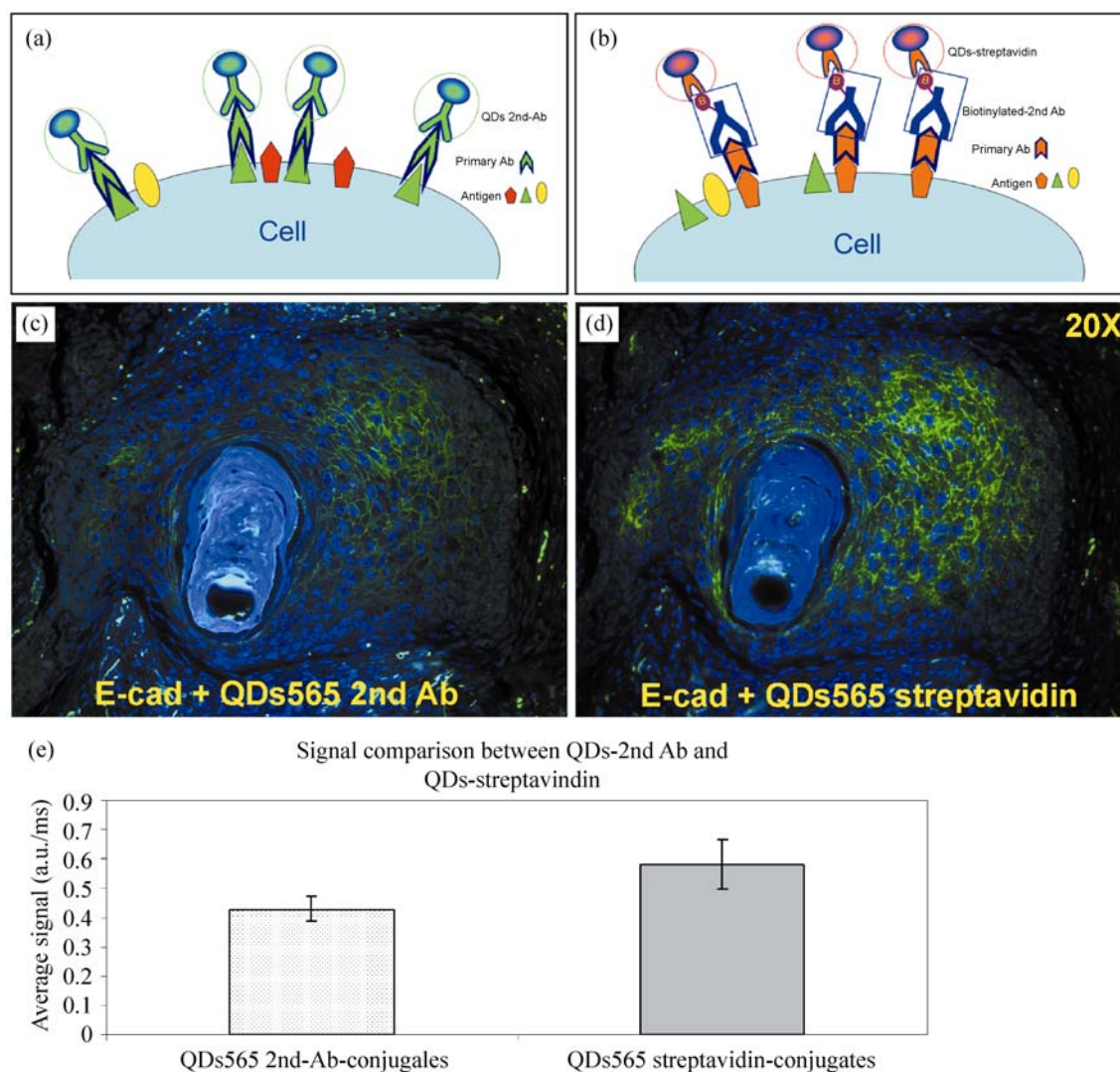


Figure 1 Comparison of single QD-IHF staining using QD-2nd Ab with QD-streptavidin: (a) cartoon showing single QD-IHF staining with QD-2nd Ab conjugates; (b) cartoon showing single QD-IHF staining with QD-streptavidin conjugates; (c) RGB image of E-cad QD-IHF staining with QD565-2nd Ab; (d) RGB image of E-cad QD-IHF staining with QD565-streptavidin; (e) signal intensity comparison between QD-2nd Ab and QD-streptavidin

For QD-IHF staining with QD-streptavidin (shown in the cartoon in Fig. 1(b)), slides were prepared as above. After the primary antibody incubation, slides were incubated with biotinylated 2nd Ab for 20 min at room temperature (RT), and washed three times with PBS (5 min each). Slides were incubated with QD565-, QD605-, or QD655-streptavidin (1:100) in 6% BSA for 1 h at 37 °C and washed three times with PBS (5 min each). After nuclei counterstaining and mounting, the slides were kept in the dark at 4 °C for visualizing and quantifying. Mouse, rabbit or goat

IgG was used as a negative control.

For sequential QD-IHF staining with QD-streptavidin, after the first primary antibody E-cad incubation (1:2000 dilution), the slides were incubated with the biotinylated 2nd Ab for 20 min at room temperature and washed three times with PBS (5 min each). Slides were then incubated with QD565-streptavidin (1:100) in 6% BSA for 1 h at 37 °C and washed three times with PBS (5 min each). After staining the first biomarker with QDs, the staining procedure was repeated from the blocking step, except the primary antibody and

QD conjugate were replaced with second biomarker EGFR (1:150) and QD605-streptavidin (1:100), respectively. Then the slides were mounted after nuclear counterstaining. For QD signal comparison, we also switched the staining sequence from EGFR with QD565-streptavidin staining as the first step to E-cad with QD605-streptavidin staining as the second. Mouse and rabbit IgG were used as a negative control.

For the cocktail staining method, we chose primary antibodies of distinct species origins, including mouse anti-human E-cad, rabbit anti-human EGFR, and goat anti-human β -cat. Therefore, for QD-2nd Abs, we selected QD565 goat F(ab')₂ anti-mouse IgG, QD605 goat F(ab')₂ anti-rabbit IgG, and QD655 rabbit F(ab')₂ anti-goat IgG, respectively. After preparation steps, the slides were incubated with the three primary antibodies against E-cad (1:2000), EGFR (1:150), and β -cat (1:2000) simultaneously for 1 h at 37 °C. After washing three times with PBS, the three QD-2nd Abs in a cocktail solution at 1:100 dilution were added to the slides with further incubation for 1 h at 37 °C. Slides were washed three times in PBS, counterstained, mounted, and stored as described above. For the sequential method (shown in the cartoon in Fig. 4(a)(ii)), the additional biomarker β -cat was stained by incubation with QD655-streptavidin following staining for E-cad with QD565-streptavidin and EGFR with QD605-streptavidin as above. The IgG with the same host species as the 2nd Ab was used as a negative control.

An Olympus microscope IX71 with CRi Nuance spectral imaging and quantifying system (CRi Inc., Woburn, MA, USA) was used to observe and quantify the QD signals. All cubed image files were collected from the FFPE tissue slides at 10-nm wavelength intervals from 500 to 800 nm with an auto exposure time at 200 \times magnification. Taking the cube with a long wavelength bandpass filter allowed transmission of all emission wavelengths above 450 nm. Both mixed and separated QD images were established after determining the QD spectral library and unmixing the cube. Background and auto-fluorescence were removed for accurate quantification of each QD signal. For comparison of the QD signals, we defined the measurement threshold as the same in each case.

An arbitrary unit (a.u.) was defined as the average fluorescence signal intensity per exposure time (ms), in which the exposure time was optimized so that the differences in fluorescence intensity among QDs were minimized and the same exposure time could be used for all QDs. These values were obtained directly from the Nuance software. Ten randomly selected fields in each sample slide were used for quantification. Data are presented as a mean of ten readings with the standard deviation (S. D.).

2. Results and discussion

We used the quantification results to evaluate the working conditions. It was found that (1) the same antigen retrieval method as used in IHC also performed well in QD-IHF staining of FFPE samples; (2) the optimized working conditions for primary antibodies in IHC also worked well for QD-IHF; (3) incubation of the QD-conjugates from Invitrogen at 10–20 nmol/L and 37 °C for 1 h was sufficient to reach a balance of the maximum staining effect with minimized non-specific binding. Non-specific binding increased when increasing either the concentration or the incubation time of the QD-conjugates, suggesting that 10–20 nmol/L of QD-conjugate may be a saturating level for staining. There was almost no significant difference in the intensity of QD signal when the concentration of QD-conjugates reached 20 nmol/L, but the non-specific binding increased directly (data not shown), suggesting that the QD binding was saturated at 20 nmol/L; (4) multiple PBS washing up to three times did not reduce the QD signal intensity. The effects of other washing buffers, such as PBS with Tween-20 (PBS-T) or Tris-buffered saline with Tween-20 (TBS-T), were similar to that of PBS. The antigen retrieval method, dilution, and incubation condition of the antibody are the main factors that affect the results of immunostaining FFPE tissues. Most researchers use the same retrieval method and incubation conditions for primary antibodies when conducting IHC and IHF stained with QDs. There are several issues that we have to address before immunostaining with QD-bioconjugates: (1) do the optimized working conditions for IHC work well for

QD-IHF? (2) how can we control the dilution ratio for QD-conjugates and the incubation conditions to obtain a balance between an optimal signal and minimized non-specific binding? (3) how can we optimize the QD-IHF staining procedure, especially in multiple staining? These potential problems have been addressed using the quantification data in our studies.

Figure 1 shows that the signal when staining with QD-2nd Ab was lower than that with QD-streptavidin (Figs. 1(c), 1(d)). The quantification results also showed that the average intensity from QD-streptavidin staining was 1.36–1.73-fold greater than that from QD-2nd Ab staining (Fig. 1(e)). The staining with QD-streptavidin had some amplification effect.

For multiplex QD staining, the sequential staining method is used by most researchers [8, 12]. To investigate whether the intensity of the QD signal after the first step changes or not after the subsequent biomarker staining and many washing steps, we initially tested sequential QD-IHF staining of E-cad with QD565-streptavidin followed by EGFR with QD605-streptavidin, and then altered this sequence. The staining signals from the two experiments were

quantified and compared. It was found that the QD intensity of E-cad staining when stained first was 0.104 ± 0.050 compared with 0.534 ± 0.132 when stained second (Fig. 2(a)). Similarly, the intensity of EGFR staining when stained first was 0.189 ± 0.104 compared with 0.565 ± 0.098 when stained second (Fig. 2(b)). This result indicated that the intensity of the first QD signal after the initial biomarker staining was reduced following the second blocking and washing steps.

In order to achieve the best staining of each biomarker using the QD-IHF sequential method, the selection of QDs may be considered. Our study has demonstrated that the intensity of QDs from Invitrogen at the same concentration was different and varied in the order QD655 > QD605 > QD565 (Fig. 3). In theory, the QD with highest intensity is recommended to be used in the first step to offset the decreasing signal when staining with QD-IHF in a sequential manner.

In order to avoid the decrease in signal observed with sequential staining, we applied three mixed primary antibodies with distinct species origins to the tissue slides and then incubated the relevant QD-2nd

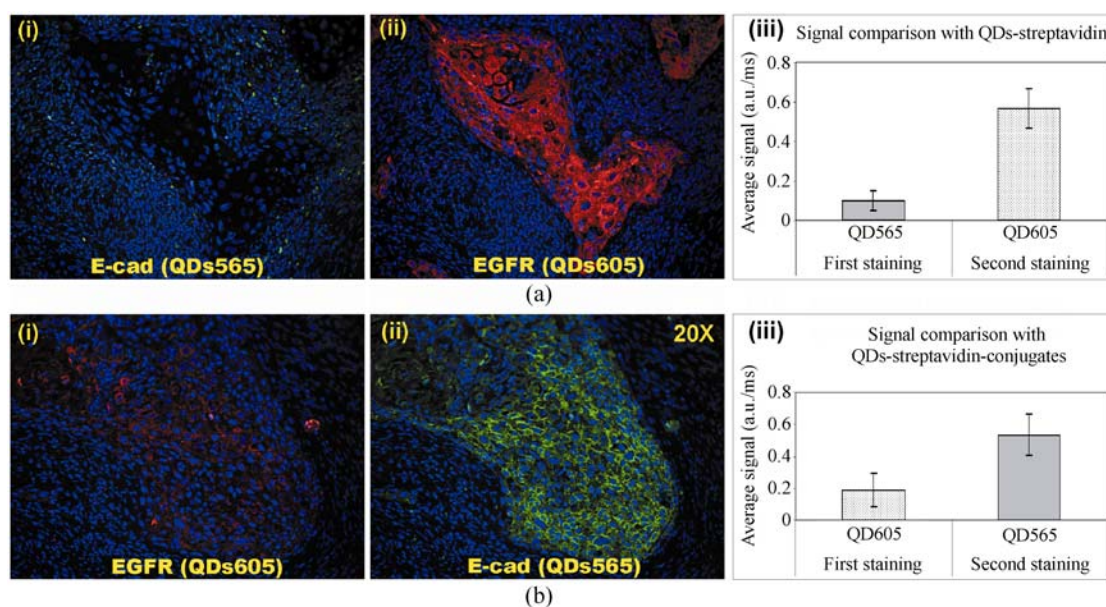


Figure 2 Comparison of the first signal with the second signal in a sequential QD-IHF staining. (a) E-cad with QD565-streptavidin as the first biomarker and EGFR with QD605-streptavidin as the second. (b) EGFR with QD605-streptavidin as the first biomarker and E-cad with QD565-streptavidin as the second. (i) unmixed first signal; (ii) unmixed second signal; (iii) quantification comparison between these two biomarkers

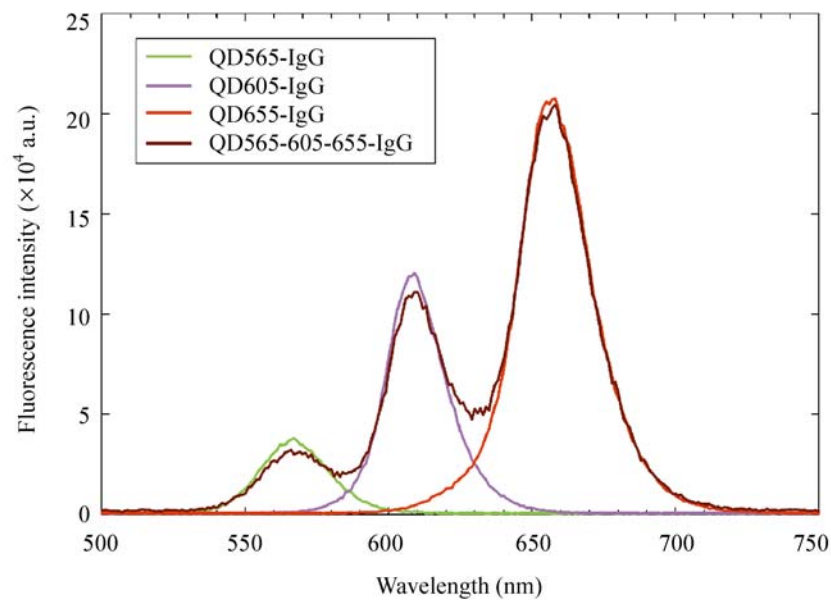


Figure 3 Comparison of QD emission intensity in single QD and cocktail PBS solutions. Fluorescence intensity of each of the three QDs was detected by QuantaMaster™ UV VIS, (Photon Technology International, Birmingham, NJ)

Abs in a cocktail solution in order to make the IHF staining efficient and simple [shown in the cartoon in Fig. 4(a)(i)]. The level of each QD signal obtained from the cocktail method was quantified and compared to that obtained with the sequential method. It was found that all of the QD signals obtained by the QD-IHF cocktail method were consistent (Fig. 4(b)). The intensities of E-cad, EGFR, and β -cat were 0.318 ± 0.015 , 0.309 ± 0.034 , and 0.362 ± 0.036 , respectively (Fig. 4(d)). In contrast, the signals from the sequential staining method were not consistent (Fig. 4(c)). Intensities of the second and the third signals were, respectively, 1.57–2.20- and 5.80–8.24-fold higher than the first signal (Fig. 4(d)).

Because the properties of nanocrystals are highly dependent on the surface environment, whether the stability with respect to the optical emission peak maximum and color purity of the QDs in such a cocktail solution may be changed is always a consideration. As recommended by the QD manufacturer, Invitrogen Cooperation, we diluted the three QDs with 6% BSA in PBS solution, and tested the signal intensity of the QDs either singly or in a cocktail solution using a spectrofluorimeter [QuantaMaster™ UV VIS, Photon Technology

International (PTI)]. Our study confirmed that the QD signals in PBS appeared at the expected wavelength with reasonable sensitivity (Fig. 3). Furthermore, the fluorescence intensity of each QD was not altered in the cocktail solution when compared to the single QD solution (Fig. 3).

One of the drawbacks of the cocktail method is that it can be challenging to find more than four primary antibodies with distinct species origins for simultaneous IHF staining, which limits the use of this method to not more than four biomarkers. In the case of multiplexing more than four biomarkers, the cocktail plus the sequential method may be applied.

3. Conclusions

We have demonstrated that the signal intensities using the QD-streptavidin-based staining method were higher than those with QD-2nd Ab. QD staining signals using the cocktail method were more consistent and stable than those obtained using the sequential method. In order to achieve the optimal signal for each biomarker in a QD-IHF multiplexed staining procedure, the staining method selection and QD intensity should be considered.

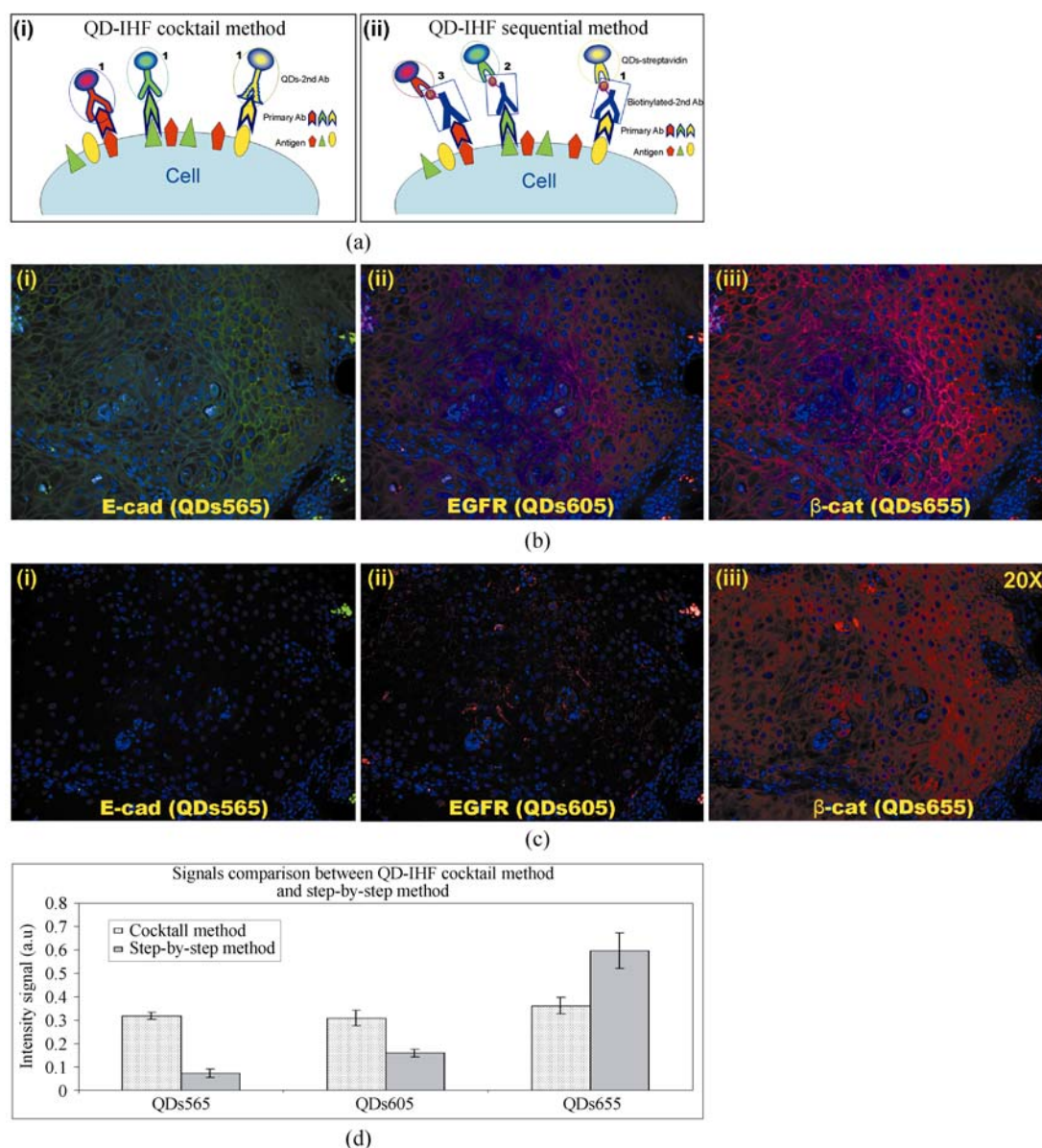


Figure 4 Comparison of the QD-IHF cocktail method with the sequential method. (a) (i) Cartoon showing cocktail QD-IHF staining with QD-2nd Ab conjugates, “1” illustrates the addition of different QDs simultaneously; (ii) cartoon showing sequential QD-IHF staining with QD-streptavidin conjugates, “1, 2, 3” represent the addition of QDs at different steps. (b) Cocktail QD-IHF staining of E-cad + EGFR + β -cat with QD565 + 605 + 655 2nd Ab-conjugates. (c) Sequential QD-IHF staining of E-cad, EGFR, and β -cat with QD565-, 605-, and 655-streptavidin conjugates, respectively: (i) unmixed E-cad (QD565) signal; (ii) unmixed EGFR (QD605) signal; (iii) unmixed β -cat (QD655) signal. (d) Quantified signal comparison between these two methods

Acknowledgements

This study was supported by DOD Grant W81XWH-07-1-0306 Project 5 and GCC Distinguished Scholar Award to Z(G). C. We thank Dr. Ximei Qian for testing QD fluorescence intensity and Dr. Anthea Hammond for her critical reading of the manuscript.

References

- [1] Ferrari, M. Cancer nanotechnology: Opportunities and challenges. *Nat. Rev. Cancer* 2005, 5, 161–171.
- [2] Nie, S. M.; Xing, Y.; Kim, G. J.; Simons, J. W. Nanotechnology applications in cancer. *Ann. Rev. Biomed. Eng.* 2007, 9, 257–288.



- [3] Smith, A. M.; Gao, X. H.; Nie, S. M. Quantum dot nanocrystals for *in vivo* molecular and cellular imaging. *Photochem. Photobiol.* 2004, *80*, 377–385.
- [4] Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly. *Nature* 2000, *405*, 665–668.
- [5] Chan, W. C. W.; Nie, S. M. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science* 1998, *281*, 2016–2018.
- [6] Wu, X. Y.; Liu, H. J.; Liu, J. Q.; Haley, K. N.; Treadway, J. A.; Larson, J. P.; Ge, N. F.; Peale, F.; Bruchez, M. P. Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nat. Biotechnol.* 2003, *21*, 41–46.
- [7] Gao, X. H. Multifunctional quantum dots for cellular and molecular imaging. *Conf. Proc. IEEE Eng. Med. Bio. Soc.* 2007, *2007*, 524–525.
- [8] Sweeney, E.; Ward, T. H.; Gray, N.; Womack, C.; Jayson, G.; Hughes, A.; Dive, C.; Byers, R. Quantitative multiplexed quantum dot immunohistochemistry. *Biochem. Biophys. Res. Commun.* 2008, *374*, 181–186.
- [9] Xing, Y.; Chaudry, Q.; Shen, C.; Kong, K. Y.; Zhau, H. E.; Chung, L. W.; Petros, J. A.; O'Regan, R. M.; Yezhelyev, M. V.; Simons, J. W.; Wang, M. D.; Nie, S. M. Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. *Nat. Protoc.* 2007, *2*, 1152–1165.
- [10] Fountaine, T. J.; Wincovitch, S. M.; Geho, D. H.; Garfield, S. H.; Pittaluga, S. Multispectral imaging of clinically relevant cellular targets in tonsil and lymphoid tissue using semiconductor quantum dots. *Modern Pathol.* 2006, *19*, 1181–1191.
- [11] True, L. D.; Gao, X. H. Quantum dots for molecular pathology: Their time has arrived. *J. Mol. Diagn.* 2007, *9*, 7–11.
- [12] Shi, C.; Zhou, G.; Zhu, Y.; Su, Y.; Cheng, T.; Zhau, H. E.; Chung, L. W. K. Quantum dots-based multiplexed immunohistochemistry of protein expression in human prostate cancer cells. *Eur. J. Histochem.* 2008, *52*, 127–133.
- [13] Yang, D. Z.; Chen, Q. F.; Wang, W. X.; Xu, S. K. Direct and indirect immunolabelling of HeLa cells with quantum dots. *Luminescence* 2008, *23*, 169–174.
- [14] Resch-Genger, U.; Grabolle, M.; Cavaliere-Jaricot, S.; Nitschke, R.; Nann, T. Quantum dots versus organic dyes as fluorescent labels. *Nat. Methods* 2008, *5*, 763–775.
- [15] Tholouli, E.; Sweeney, E.; Barrow, E.; Clay, V.; Hoyland, J. A.; Byers, R. J. Quantum dots light up pathology. *J. Pathol.* 2008, *216*, 275–285.
- [16] Muller, S.; Su, L.; Tighiouart, M.; Saba, N.; Zhang, H. Z.; Shin, D. M.; Chen, Z. G. Distinctive E-cadherin and epidermal growth factor receptor expression in metastatic and nonmetastatic head and neck squamous cell carcinoma: Predictive and prognostic correlation. *Cancer* 2008, *113*, 97–107.

Applying E_{max} model and bivariate thin plate splines to assess drug interactions

Maiying Kong¹ & J. Jack Lee²

¹Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, Kentucky 40292, U.S.A., ²Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Unit 1411, 1515 Holcombe Boulevard, Houston, Texas 77030, U.S.A.

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Statistical methods
 - 3.1. Estimating dose-effect curves
 - 3.2. Predicting additive effects
 - 3.3. Assessing drug interactions using bivariate thin plate splines
4. Case studies
 - 4.1. Case study 1: cancer cells grown in a medium with 2.3 μ M folic acid (low FA experiment)
 - 4.2. Case study 2: cancer cells grown in a medium with 78 μ M folic acid (high FA experiment)
5. Summary and perspective
6. Acknowledgement
7. References

1. ABSTRACT

We review the semiparametric approach previously proposed by Kong and Lee and extend it to a case in which the dose-effect curves follow the E_{max} model instead of the median effect equation. When the maximum effects for the investigated drugs are different, we provide a procedure to obtain the additive effect based on the Loewe additivity model. Then, we apply a bivariate thin plate spline approach to estimate the effect beyond additivity along with its 95% point-wise confidence interval as well as its 95% simultaneous confidence interval for any combination dose. Thus, synergy, additivity, and antagonism can be identified. The advantages of the method are that it provides an overall assessment of the combination effect on the entire two-dimensional dose space spanned by the experimental doses, and it enables us to identify complex patterns of drug interaction in combination studies. In addition, this approach is robust to outliers. To illustrate this procedure, we analyzed data from two case studies.

2. INTRODUCTION

Studies of interactions among biologically active agents, such as drugs, carcinogens, or environmental pollutants have become increasingly important in many branches of biomedical research. For example, in cancer chemotherapy, the therapeutic effect of many anticancer drugs is limited when they are used as single drugs. Finding combination therapies with increased treatment effect and decreased toxicity is an active and promising research area (1). An effective and accurate evaluation of drug interaction for *in vitro* and/or *in vivo* studies can help to determine whether a combination therapy should be further investigated.

The literature supports the Loewe additivity model as the gold standard for defining drug interactions (2-5). The Loewe additivity model defines an additive effect based on the following equation

$$\frac{d_1}{D_{y,1}} + \frac{d_2}{D_{y,2}} = 1. \quad (E1)$$

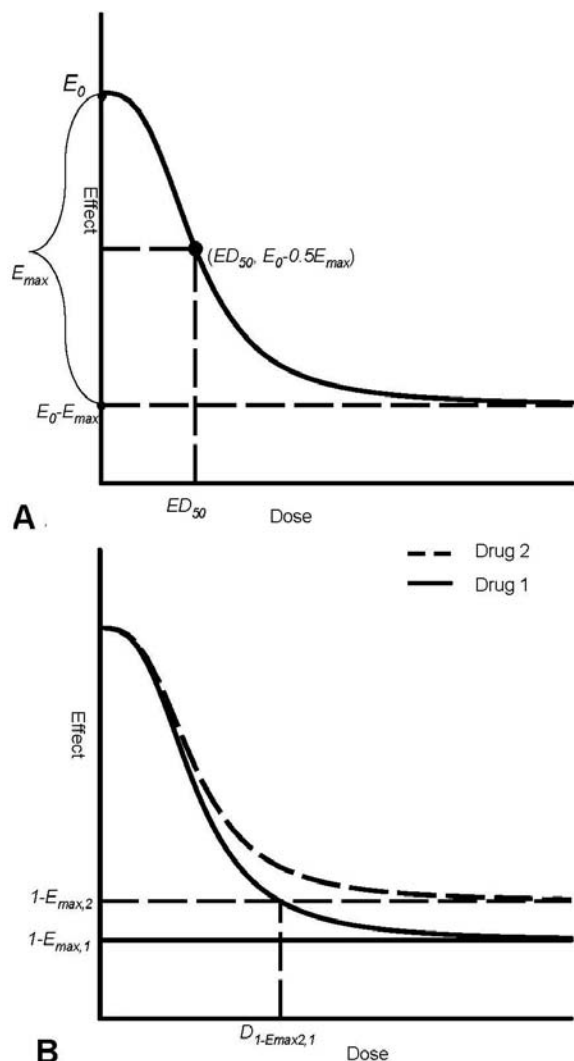


Figure 1. Dose-effect curves. Panel A shows a typical curve with the maximum effect, E_{max} , less than 1. ED_{50} is the dose required to produce half of the maximum effect, $E_0 - 0.5E_{max}$. Panel B shows two dose-effect curves with different maximum effects, say, $E_{max,1} > E_{max,2}$. In Panel B, drug 1 at dose level $D_{1-E_{max,2,1}}$ produces the maximum effect produced by drug 2 alone.

Here y is the predicted additive effect, which is produced by the combination dose (d_1, d_2) when the two drugs do not interact; and $D_{y,1}$ and $D_{y,2}$ are the respective doses of drug 1 and drug 2 required to produce the same effect y when applied alone. If we know the dose-effect relationship for each single agent, say $E(d) = f_i(d)$ for agent i ($i=1,2$), we are able to obtain the dose $D_{y,i}$ by using the inverse function of f_i , denoted as $f_i^{-1}(y)$. By replacing $D_{y,1}$ and $D_{y,2}$ in equation (E1) with $f_1^{-1}(y)$ and $f_2^{-1}(y)$, respectively, we can obtain an equation that includes the single variable y , i.e.,

$$\frac{d_1}{f_1^{-1}(y)} + \frac{d_2}{f_2^{-1}(y)} = 1. \quad (E2)$$

By solving equation (E2), we can obtain the predicted additive effect y . If the observed effect at (d_1, d_2) is more than (equal to, or less than) the predicted effect, we say that the combination dose (d_1, d_2) is synergistic (additive, or antagonistic).

In our previous studies (6-8), we found that Chou and Talalay's (9) median effect equation was appropriate to describe the dose-effect relationships. Chou and Talalay's median effect equation, in its nonlinear form, can be written as follows:

$$E = \frac{(d / ED_{50})^m}{1 + (d / ED_{50})^m} \quad (E3)$$

where ED_{50} is the dose required to produce 50% of the maximum effect, and m is the slope factor (Hill coefficient), measuring the sensitivity of the effect to the dose range of the drug. For the data in the case studies (see Section 4 for details), we found that the median effect equation (E3) could not adequately describe the marginal dose-effect relationship because the plateau of the effect does not go to zero when a large dose level of a drug is applied. Instead, the following E_{max} model (E4) presented by Ting (10) describes the dose-effect relationship very well:

$$E = E_0 - \frac{(d / ED_{50})^m E_{max}}{1 + (d / ED_{50})^m}. \quad (E4)$$

In the E_{max} model (E4), E_0 is the base effect, corresponding to the measurement of response when no drug is applied; E_{max} is the maximum effect attributable to the drug; ED_{50} is the dose level producing half of E_{max} , i.e., ED_{50} is the dose level required to produce the effect at a value of $E_0 - 0.5E_{max}$ (Figure 1.A); d is the dose level, which produces the effect E . Thus, $E_0 - E_{max}$ will be the asymptotic net effect when a large dose of the drug is applied. Different maximum effects for agents may reflect different mechanisms of action for the drugs (11). For *in vitro* studies, cell growth is commonly used as an endpoint to measure the effect of inhibitors. When no drugs (or, no inhibitors) are applied, the cell proliferation obtains its largest value. In this case, the dose-effect curve is similar to the one shown in Figure 1.A, where $E_{max} > 0$. The effect range determined by the dose-effect curve lies between $(E_0 - E_{max}, E_0)$, and the asymptotic measurement for the maximum drug effect is $E_0 - E_{max}$.

In the investigation of drug interactions, theoretically, we expect the measurements for the endpoints to be similar when no drug is applied. We use the measurements that are made without any drugs applied as controls. However, we realize that environmental factors other than the experimental conditions may lead to different measurements for the controls under different environments. Thus, we may need to standardize the observed effects by the mean of the control for each environmental condition (1, 6), and then take $E_0 = 1$. In this paper, we consider the following dose-effect curve for each drug:

$$E = 1 - \frac{(d / ED_{50})^m E_{max}}{1 + (d / ED_{50})^m}, \quad (E5)$$

which assumes an effect at value 1 when no drug is applied. Once we obtain the dose-effect curve for each single drug, we can use the Loewe additivity model (E1) to obtain the additive effect for any combination dose, particularly, for the combination doses with observed effects. Thus, we may obtain the differences in observed effects and the predicted additive effect at each observed combination dose. We use the bivariate thin plate splines approach (12) to estimate the relationship between these differences and the combination doses. Consequently, we obtain a response surface of the differences over the combination doses, and can construct 95% confidence surfaces of the response surface. When the dose-response curves decrease with increasing dose, an observed effect that is smaller in magnitude than the prediction of Loewe additivity implies that the observed effect is stronger than the predicted effect, indicating that the combination dose is synergistic. Conversely, an observed effect that is larger in magnitude than the prediction of Loewe additivity implies that the observed effect is weaker than the predicted effect, indicating that the combination dose is antagonistic. However, these inferences should be made based on sound statistical considerations. Based on the fitted response surface and its upper and lower confidence surfaces, we can judge whether the difference is significantly less than zero, not different from zero, or greater than zero, and we can determine the patterns of drug interaction in terms of synergy, additivity, and antagonism. This paper is organized as follows. In Section 3.1, we describe the underlying stochastic assumption for the dose-effect curve and the procedure to estimate the parameters in each marginal dose-effect curve. In Section 3.2, we explain how we obtained the additive response surface based on the Loewe additivity model, in particular for studies in which the maximum effects of the drugs are different. In Section 3.3, we explain how we assessed the response surface beyond the additivity surface and how we constructed its 95% confidence surfaces. These procedures allow us to identify drug interactions in terms of synergy, additivity, or antagonism for all of the combination doses in the region containing the combination design points. In Section 4, we illustrate the procedure introduced in Section 3 by analyzing real data in two case studies. The last section is devoted to a short summary and perspective.

3. STATISTICAL METHODS

Assume that the observed data are (d_{1i}, d_{2i}, E_i) for $i=1, \dots, n$. For each i , (d_{1i}, d_{2i}) is the observed combination dose and E_i is the corresponding observed effect. When only a single drug is applied (drug 1 or drug 2), we refer to the observations as marginal observations. That is, the marginal observations for drug 1 are the observations (d_{1i}, d_{2i}, E_i) with $d_{2i}=0$ ($i=1, \dots, n$), and the marginal observations for drug 2 are the observations (d_{1i}, d_{2i}, E_i) with $d_{1i}=0$ ($i=1, \dots, n$). The marginal dose-effect curves are estimated based on the marginal observations, which we present in Section 3.1. It is commonly accepted that the additive effect should be obtained based on the dose-effect relationships for each individual drug. In Section 3.2, we explain how we obtained the predicted

effect at combination dose (d_1, d_2) based on the Loewe additivity model (E1) and the marginal dose-effect curves (E5). We denote the predicted effect as $\hat{F}_p(d_1, d_2)$. By definition, there is no drug interaction when only a single drug is applied. Therefore, the term for drug interaction is meaningful only for the combination dose (d_1, d_2) with nonzero d_1 and d_2 . In Section 3.3, we develop a procedure to estimate the effect beyond additivity for any combination dose (d_1, d_2) with nonzero d_1 and d_2 , denoted by $\hat{f}(d_1, d_2)$.

3.1. Estimating dose-effect curves

Chou and Talalay (9), Chou (4), and Kong and Lee (6) estimated the parameters in the median effect equation (E3) by using the transformation $\log E/(1-E) = m \log(d/ED_{50}) = \alpha + m \log(d)$ and applying the least squares method in the linear regression setting, where $\alpha = -m \log(ED_{50})$. The case studies we evaluated (see Section 4) included many low doses, the effects of which are larger than 1 after setting the effect at the control level to be 1. Thus, a similar transformation for models (E3) and (E5) cannot be carried out. Because the measurements are continuous, we propose applying nonlinear least squares regression to estimate the parameters in models (E3) and (E5) with the assumption that a stochastic error with $N(0, \sigma^2)$ exists on the right-hand side of the two models. We note that estimating the dose-effect curve for drug i requires only the marginal observations for drug i with $i=1, 2$. We apply nonlinear least squares regression to estimate the parameters in the marginal dose-effect curves in the two case studies (shown in Section 4).

3.2. Predicting additive effects

We obtain the predicted effect based on the Loewe additivity model (E1) when model (E5) is applied as the marginal dose-effect curve for each drug. When model (E5) is applied, the dose required to produce effect E is given by

$$d = ED_{50} \left(\frac{1-E}{E - (1-E_{max})} \right)^{1/m}.$$

However, the maximum effects for the two drugs may be different. Without a loss of generality, we assume that the maximum effect of drug 1 is larger than the maximum effect of drug 2, i.e., $E_{max,1} > E_{max,2}$. For this case, when the dose-effect curves are decreasing, neither drug applied alone can produce an effect in $(0, 1 - E_{max,1})$ (Figure 1.B). Based on the Loewe additivity model (E1), we can see that the predicted effect will be in the interval of $(1 - E_{max,1}, 1)$ for any combination dose (d_1, d_2) .

Recall that the Loewe additivity model (E1) can be rewritten as $d_1 + (D_{y,1}/D_{y,2})d_2 = D_{y,1}$, and the ratio $D_{y,1}/D_{y,2}$ (denoted as $\rho(y)$), is often called the relative potency of drug 2 versus drug 1 at effect level y , which means that the effect of 1 unit of drug 2 produces the same effect as $\rho(y)$ units of drug 1. Generally speaking, the relative potency $\rho(y)$ is dose-dependent (7). When there is no drug interaction, the effect of the combination dose (d_1, d_2) produces the same effect as drug 1 alone at dose level

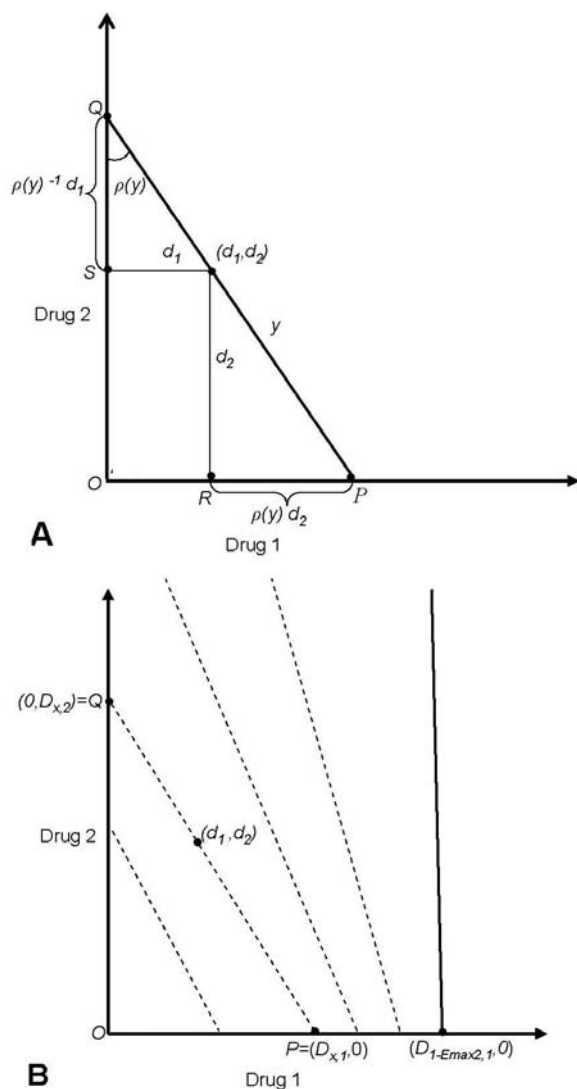


Figure 2. Additive isoboles. Panel A shows additive isobole under the Loewe additivity model. Any combination dose (d_1, d_2) on the line \overline{PQ} produces the same effect as drug 1 alone at dose $D_{y,1}$ ($d_1 + \rho(y)d_2$), or drug 2 alone at dose $D_{y,2}$ ($\rho(y)^{-1}d_1 + d_2$), y is the predicted effect for any combination dose at the line \overline{PQ} , and $\rho(y)$ is the relative potency at the effect level y . Panel B shows that the additive isoboles associated with the effect level in $(1-E_{max,2}, 1)$ cover the bound between the two solid vertical lines under the assumption $E_{max,1} > E_{max,2}$. Each dashed line corresponds to an isobole.

$D_{y,1}$, which equals $d_1 + \rho(y)d_2$, or drug 2 alone at dose $D_{y,2}$, which equals $\rho(y)^{-1}d_1 + d_2$ (Figure 2.A). All the combination doses (d_1, d_2) on the line \overline{PQ} have the predicted effect y , where \overline{PQ} is the line connecting the points $P=(D_{y,1}, 0)$ and $Q=(0, D_{y,2})$ (Figure 2.A). This line \overline{PQ} is often called an additive isobole (3, 4).

When $E_{max,1} > E_{max,2}$, as illustrated in Figure 1.B, we can calculate the dose of drug 1 required to produce the maximum effect of drug 2, i.e.,

$$D_{1-E_{max,2},1} = ED_{50,1} \left(\frac{E_{max,2}}{E_{max,1} - E_{max,2}} \right)^{1/m_1}.$$

Note that the range of the effect for drug 2 is $(1-E_{max,2}, 1)$, which could be produced by drug 1 alone at a dose level between 0 and $D_{1-E_{max,2},1}$. Based on the Loewe additivity model, for any

level of effect y in $(1-E_{max,2}, 1)$, the associated additive isobole is the line connecting $(D_{y,1}, 0)$ and $(0, D_{y,2})$. When y varies from $1-E_{max,2}$ to 1, the dose of drug 1 required to produce effect y varies from $D_{1-E_{max,2},1}$ to 0, while the dose

of drug 2 required to produce effect y varies from infinitely large to 0. In particular, when y is close to $1-E_{max,2}$, the dose of drug 1 required to produce effect y is close to $D_{1-E_{max,2},1}$, and the dose of drug 2 required to produce

effect y goes to infinity. Figure 2.B shows four typical additive isoboles (dashed lines), which connect equally effective doses of drug 1 and drug 2 at different effect levels. From left to right, the effect level decreases in magnitude. The additive isoboles may not be parallel because the relative potency may not be constant. When y varies in $(1-E_{max,2}, 1)$, all the additive isoboles cover the region between the two solid vertical lines (Figure 2.B). Meanwhile, any combination dose (d_1, d_2) with $d_1 < D_{1-E_{max,2},1}$ must lie on one of these isoboles. Therefore,

for any combination dose (d_1, d_2) with $d_1 < D_{1-E_{max,2},1}$, the predicted additive effect, say y , can be obtained by solving the following nonlinear equation for E :

$$\frac{d_1}{ED_{50,1} \left(\frac{1-E}{E - (1-E_{max,1})} \right)^{1/m_1}} + \frac{d_2}{ED_{50,2} \left(\frac{1-E}{E - (1-E_{max,2})} \right)^{1/m_2}} = 1.$$

Now we examine the predicted effect for the combination dose (d_1, d_2) with $d_1 \geq D_{1-E_{max,2},1}$. When $d_1 \geq D_{1-E_{max,2},1}$, drug 1 alone at dose d_1 produces an effect

$$E = 1 - \frac{(d_1 / ED_{50,1})^{m_1} E_{max,1}}{1 + (d_1 / ED_{50,1})^{m_1}},$$

$1-E_{max,2}$, which cannot be produced by drug 2 alone at any dose level. In this case, if the effect of the combination dose is more than the effect produced by drug 1 alone, then drug 2 potentiates the effect of drug 1, and synergy occurs because the predicted additive effect is the effect produced by drug 1 alone at dose level d_1 . Alternatively, because drug 2 alone cannot produce such an effect, we could consider $D_{y,2}$ to be infinitely large. Thus, the Loewe additivity model is reduced to $d_1/D_{y,1}=1$, and the predicted additive effect is the effect produced by drug 1 alone at $D_{y,1}$. No matter which approach we take, the predicted effect y is the same. We can determine the predicted effect y from the following equation:

$$y = f_1(D_{y,1}) = f_1(d_1) = 1 - \frac{(d_1 / ED_{50,1})^{m_1} E_{max,1}}{1 + (d_1 / ED_{50,1})^{m_1}}.$$

Thus, we can obtain the predicted effect for any combination dose (d_1, d_2) . Using notation similar to that

from the previous study of Kong and Lee (6), we denote the predicted effect at the combination dose (d_1, d_2) as $\hat{F}_p(d_1, d_2)$. In the following subsection, we develop a procedure to estimate the effect beyond additivity, denoted by $\hat{f}(d_1, d_2)$ and to construct its 95% point-wise confidence interval and simultaneous confidence interval. We assess the drug-drug interaction based on the estimated $\hat{f}(d_1, d_2)$ and its confidence intervals.

3.3. Assessing drug interactions using bivariate thin plate splines

By definition, there is no drug interaction when a single drug is used alone. Therefore, we set the differences between the observed and predicted effects at zero for the marginal observations, that is, for the combination doses (d_1, d_2) with only one nonzero component. We apply a bivariate thin plate spline to estimate the differences as a function of the combination dose, say, $f(d_1, d_2)$. When the dose-effect curves are decreasing, $f(d_1, d_2) < 0$ indicates that the observed effect is more than the predicted effect at (d_1, d_2) , thus the combination dose (d_1, d_2) is synergistic. Inversely, $f(d_1, d_2) > 0$ indicates that the combination dose (d_1, d_2) is antagonistic. Kong and Lee (6) used the different observed combination doses as the knots for the bivariate thin plate splines (12). The choice of knots works well when the number of combination doses is not large and the combination doses are not close, such as those from factorial designs or uniform designs (13). However, when ray designs are applied and the doses are low, the combination doses are very close and some columns of the design matrix (i.e., Ω and Z_1 in the following notations) may be highly correlated, which results in a nearly singular matrix for estimating the parameters in the function f . If that happens, a low rank smoothing thin plate spline (14) should be applied to avoid the singularity of the involved matrix due to the low rank of the design matrix. An example of such a low rank smoothing thin plate spline is the knots formed by selecting the observed combination doses with distances larger than some pre-specified small number. Alternatively, one may use an appropriate transformation of the dose, such as the log-transformation, to evenly distribute the experimental combination doses in certain regions in order to improve the ability to estimate the effect beyond additivity using bivariate thin plate splines.

Suppose the selected knots are $(\kappa_{1k}, \kappa_{2k})$ ($k=1, \dots, K$), then the bivariate thin plate spline can be expressed by the following form:

$$f(d_1, d_2) = \gamma_0 + \gamma_1 d_1 + \gamma_2 d_2 + \sum_{k=1}^K v_k \eta \left(\left\| (d_1, d_2)^T - (\kappa_{1k}, \kappa_{2k})^T \right\| \right),$$

where $\gamma = (\gamma_0, \gamma_1, \gamma_2)^T$ and $v = (v_1, \dots, v_K)^T$ are the parameters in the thin plate spline function f , and

$$\eta(r) = \frac{1}{16\pi} r^2 \log r^2 \text{ for } r > 0 \text{ and } \eta(r) = 0 \text{ for } r = 0. \quad \text{The}$$

distance in the expression is the Euclidean distance. Let us denote

$$\begin{aligned} \Omega &= \left[\eta \left(\left\| (\kappa_{1k}, \kappa_{2k})^T - (\kappa_{1k'}, \kappa_{2k'})^T \right\| \right) \right]_{1 \leq k, k' \leq K} \in R^{K \times K}, \\ Y_R &= \left[\left(E_1 - \hat{F}_p(d_{11}, d_{21}) \right) 1_{\{d_{11} \neq 0 \text{ \& } d_{21} \neq 0\}}, \dots, \left(E_n - \hat{F}_p(d_{1n}, d_{2n}) \right) 1_{\{d_{1n} \neq 0 \text{ \& } d_{2n} \neq 0\}} \right]^T \in R^{n \times 1}, \\ X &= [1, d_{1i}, d_{2i}]_{1 \leq i \leq n} \in R^{n \times 3}, Z_1 = \left[\eta \left(\left\| (d_{1i}, d_{2i})^T - (\kappa_{1k}, \kappa_{2k})^T \right\| \right) \right]_{1 \leq k \leq K, 1 \leq i \leq n} \in R^{n \times K}, \\ \text{and } T^T &= [1, \kappa_{1k}, \kappa_{2k}]_{1 \leq k \leq K} \in R^{K \times 3}. \end{aligned}$$

The coefficient γ and V can be obtained by minimizing the following penalized residual sum of squares:

$$(Y_R - X\gamma - Z_1 v)^T (Y_R - X\gamma - Z_1 v) + \lambda v^T \Omega v \quad (E6)$$

subject to $Tv = 0$.

Following the notation by Kong and Lee (6) and Green and Silverman (12), consider a QR decomposition of T^T , say $T^T = FG$, where F is a $K \times K$ orthogonal matrix and G is a $K \times 3$ upper triangular matrix. Let F_1 be the first three columns of F , and F_2 be the last $K-3$ columns of F . We can show that $Tv = 0$ if and only if V can be expressed as $F_2 \xi$ for some ξ . Thus, the penalized residual sum of squares can be expressed as

$$(Y_R - X\gamma - Z_1 F_2 \xi)^T (Y_R - X\gamma - Z_1 F_2 \xi) + \lambda \xi^T F_2^T \Omega F_2 \xi.$$

$$\text{Set } u = (F_2^T \Omega F_2)^{-\frac{1}{2}} \xi \quad \text{and} \quad Z = Z_1 F_2 (F_2^T \Omega F_2)^{-\frac{1}{2}},$$

where $(F_2^T \Omega F_2)^{-\frac{1}{2}}$ is the matrix square root of $F_2^T \Omega F_2$.

The penalized residual sum of squares can be expressed as

$$(Y_R - X\gamma - Zu)^T (Y_R - X\gamma - Zu) + \lambda u^T u \quad (E7)$$

Based on the approach proposed by Ruppert, Wand, and Carroll (15) and Wang (16), and detailed by Kong and Lee (6) in this setting, the parameters in terms of γ and u can be obtained by solving the following mixed effect model:

$$Y_R = X\gamma + Zu + \varepsilon \quad (E8)$$

where

$$\begin{pmatrix} u \\ \varepsilon \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 I_{K-3} & 0 \\ 0 & \sigma_\varepsilon^2 I_n \end{pmatrix} \right).$$

Thus, the parameters can be estimated by $\begin{pmatrix} \hat{\gamma} \\ \hat{u} \end{pmatrix} = (C^T C + \hat{\lambda} D)^{-1} C^T Y_R$ with $\hat{\lambda} = \hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$,

$C = [X \ Z]$, and $D = \text{diag}(0, 0, 0, 1, \dots, 1)$, where the number of zeros in the matrix D corresponds to the number of γ_i 's ($i=0, 1, 2$) and the number of ones corresponds to the number of u_i 's ($i=1, \dots, K-3$). Under these

notations, for any combination dose (d_1, d_2) , $f(d_1, d_2)$ can be predicted by $\hat{f}(d_1, d_2) = \hat{\gamma}_0 + \hat{\gamma}_1 d_1 + \hat{\gamma}_2 d_2 + Z_0 \hat{u}$ with

$$Z_0 = \left[\eta \left(\left\| (d_1, d_2)^T - (\kappa_{1k}, \kappa_{2k})^T \right\| \right) \right]_{1 \leq k \leq K} F_2 (F_2^T \Omega F_2)^{-1/2}, \text{ and}$$

an approximate $100(1-\alpha)\%$ point-wise confidence interval for $f(d_1, d_2)$ can be constructed by

$$\hat{f}(d_1, d_2) \pm z_{\alpha/2} \hat{\sigma}_\varepsilon \sqrt{C_d^T (C^T C + \hat{\lambda} D)^{-1} C_d}, \quad (E9)$$

where $C_d = (I, d_1, d_2, Z_0)$ and $z_{\alpha/2}$ is the upper $\alpha/2 \times 100\%$ percentile of the standard normal distribution. Thus, we can construct a 95% point-wise confidence band for $f = 0$ by taking the intercept lines of the lower and upper confidence surfaces with the dose plane. We then claim that the combination doses in the area outside the confidence bound with $f < 0$ are synergistic, the combination doses inside the bound are additive, and the combination doses in the area outside the bound with $f > 0$ are antagonistic.

Based on the 95% point-wise confidence intervals constructed from (E9), some combination doses that are additive may be claimed as synergistic or antagonistic. To be conservative and to control the family-wise error rate, we also construct the 95% lower and upper simultaneous confidence surfaces, which are based on a format similar

to that of equation (E9) except that $z_{\alpha/2}$ is replaced by $\sqrt{EDF \times F_{EDF, n-EDF}^{\alpha}}$ (17), where EDF is the effective degrees of freedom from the resulting bivariate smoothing splines (12) and is defined as the trace of the matrix $C(C^T C + \hat{\lambda} D)^{-1} C^T$, and $F_{EDF, n-EDF}^{\alpha}$ is the upper $100 \times \alpha$ percentile of the F distribution with EDF and $n-EDF$ degrees of freedom. Here n is the total number of observations except controls. A 95% simultaneous confidence band for $f = 0$ can be formed by taking the intercept lines of the 95% lower and upper simultaneous confidence surfaces with the dose plane. For the two case studies presented in the next section, we will present the plots of different patterns of drug interaction based on the respective 95% point-wise confidence intervals and the 95% simultaneous confidence band.

4. CASE STUDIES

The following two data sets were provided by Dr. William R. Greco. The data were collected during a study of the joint effect of trimetrexate (TMQ) and AG2034 on cells grown in medium with different concentrations of folic acid (FA): 2.3 μM in the first experiment (the low FA experiment), and 78 μM in the second experiment (the high FA experiment). TMQ is a lipophilic inhibitor of the enzyme dihydrofolate reductase, and AG2034 is an inhibitor of the enzyme glycinamide ribonucleotide formyltransferase. The unit of drug concentration is the micromole (μM) for all data analyzed in the case studies. The endpoint was the growth of human ileocecal adenocarcinoma (HCT-8) cells in 96-well assay plates as measured by the sulforhodamine B (SRB) protein stain. The drug treatments were randomly assigned to the cells in the assay wells. Each 96-well plate included 8 wells as instrumental blanks (no cells); thus 88 wells were used for drug treatments. Five replicate plates were used for each set of 88 treated wells. Each of the two large data sets were obtained from two 5-plate stacks with a maximum of 880 treated wells per experiment. Each experiment included 110 control wells, in which no drugs were applied to the cells. Ray designs were used, with the experimental doses

being distributed in 14 rays, including two rays for TMQ and AG2034 when used alone. The complete details and mechanistic implications of the study were reported by Faessel *et al* (18). Assuming that the first observation recorded in each dose or combination dose from the first 5-plate stack was from the same plate, say the 1st plate, the second observation from the 2nd plate, and so on, and also assuming that the first observation recorded in each dose or combination dose from the second 5-plate stack was from the same plate, say the 6th plate, the second observation from the 7th plate, and so on, we have a total of 10 plates for each of the two data sets.

To examine whether there is a significant difference among the plates, we applied one-way analysis of variance (ANOVA) to the controls in each data set. The p-values were 0.001 for the low FA experimental data and 0.005 for the high FA experimental data. The results indicate a significant plate effect among the 10 plates for each experiment, that is, the inter-plate variability is high. To attenuate the effect from the inter-plate variability, we applied a standardization procedure to each data set, dividing the effect readings by the mean of the controls in each associated plate. Thus, the mean for the controls within each plate was standardized to 1, and the effect for the controls was treated as 1. In addition to 110 controls for each experiment, the data included 761 observations for the low FA experiment and 769 observations for the high FA experiment. We applied the statistical method described in Section 3 to each of the two standardized data sets. We present the results for each experiment in the following two subsections.

Lee *et al* (19) performed extensive exploratory analyses on the same data sets and identified 129 outliers out of 871 (14.8%) effect readings in the low FA experiment and 126 outliers out of 879 (14.3%) effect readings in the high FA experiment. To compare our findings with the results previously obtained (19), we removed the outliers from the data and then again applied the statistical method described in Section 3. For each experiment, we report the detailed analyses of the original data set and of the modified data set that excluded the outliers.

4.1. Case study 1: cancer cells grown in a medium with 2.3 μM folic acid (low FA experiment)

In this experiment, the cells were grown in a medium with 2.3 μM folic acid. We fitted marginal dose-effect curves for TMQ and AG2034 by using both the median effect equation (E3) and the E_{\max} model (E5). The dose levels for TMQ when applied alone were 5.47×10^{-6} , 4.38×10^{-5} , 1.38×10^{-4} , 4.38×10^{-4} , 8.75×10^{-4} , 1.75×10^{-3} , 3.5×10^{-3} , 7×10^{-3} , 2.21×10^{-2} , 7×10^{-2} , and 0.56 μM , and the dose levels for AG2034 when applied alone were 2.71×10^{-5} , 2.71×10^{-4} , 6.87×10^{-4} , 2.17×10^{-3} , 4.3×10^{-3} , 8.7×10^{-3} , 1.74×10^{-2} , 3.48×10^{-2} , 0.11, 0.3475, and 2.78 μM . Note that some effect readings at low doses or combination doses are greater than 1, thus, the logit transformation could not be carried out. We applied

Table 1. Estimated parameters for the E_{max} models in the case studies

Drug name	Low FA ¹			High FA ²		
	E_{max} ³	ED_{50} ⁴	Slope m ⁵	E_{max} ³	ED_{50} ⁴	Slope m ⁵
TMQ ⁶	0.8810 (0.0161)	0.0013 (0.0001)	2.2496 (0.2330)	0.8847 (0.0326)	0.0134 (0.0015)	3.7230 (0.7323)
AG2034 ⁷	0.8688 (0.0154)	0.0060 (0.0003)	3.1644 (0.3703)	0.8184 (0.0311)	0.4700 (0.0540)	1.6869 (0.2400)

¹Estimated parameters for the marginal dose-effect curves in the low-concentration folic acid experiment; ² estimated parameters for the marginal dose-effect curves in the high-concentration folic acid experiment; ³ maximum effect attributable to the drug; ⁴ dose level producing half of E_{max} ; ⁵ factor (Hill coefficient) measuring sensitivity of the effect to the drug dose range; ⁶ trimetrexate; ⁷ experimental drug.

nonlinear least squares regression to estimate the parameters in models ($E3$) and ($E5$). Figures 3.A and 3.B show the respective fitted marginal dose-effect curves for TMQ and AG2034 with the dose levels shown on a log scale. The dotted-dashed lines are the curves based on the median effect model ($E3$), and the solid lines are the dose-effect curves based on the E_{max} model ($E5$). From the fitted dose-effect curves, we found that the E_{max} model ($E5$) provided a much better fit than the median effect equation for the marginal data. Therefore, we chose the E_{max} model ($E5$) to describe the dose-effect relationship in this case study. The parameters estimated for TMQ and AG2034 are shown in the three columns under the title “Low FA” in Table 1. Here, the estimate of $E_{max, TMQ}$ is slightly larger than the estimate of $E_{max, AG2034}$. We plotted the distribution of the combination doses using the original scale (not shown) and found that most of the combination doses were crowded in the region of the low doses, which could cause a singularity of the involved matrices due to the low rank of Ω and Z_1 used for estimating the effect beyond additivity when using bivariate thin plate splines (see Section 3.3). Hence, we applied a log transformation of the form $\log(\text{dose}+\delta)$ for each dose level, where δ is a small number, say 2.74×10^{-6} , half of the smallest dose level for the two drugs when applied alone. We plotted the distribution of the combination doses on the $\log(\text{dose}+\delta)$ scale, as shown in Figure 3.C. The points on the horizontal line in Figure 3.C are the doses of TMQ on the $\log(\text{dose}+\delta)$ scale; the points on the vertical line are the doses of AG2034 on the $\log(\text{dose}+\delta)$ scale; and the points on each of the remaining 12 design rays are the combination doses at each ray with each dose component on the $\log(\text{dose}+\delta)$ scale. The 12 design rays for combination doses appearing left to right in Figure 3.C correspond to the combination doses at 12 ratios of TMQ to AG2034, i.e., 1:250, 1:125, 1:50, 1:20, 1:10, 1:5, 1:5, 2:5, 4:5, 2:1, 5:1, 10:1. The 12 rays are denoted by the letters E, F, G, H, I, J, K, L, M, N, O, P, which represent the respective curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set for the low FA experiment. Note that rays 3 and 9, denoted by J and K, indeed have the same fixed dose ratio. To obtain the predicted additive effects, we applied the procedure described in Section 3.2, keeping the dose levels on the original scale. The contour plot of the predicted additive effect is shown in Figure 3.D. Note that the effect levels for TMQ applied alone are obtained from $(1-E_{max, TMQ}, 1)$, which is $(0.1190, 1)$, and the effect levels for AG2034 applied alone are obtained from $(1-E_{max, AG2034}, 1)$, which is $(0.1312, 1)$. The vertical line with contour level 0.13 is the predicted effect produced by TMQ alone. The plot of the differences of the observed effects and predicted effects versus the dose levels of AG2034 on $\log(\text{dose}+\delta)$

scale is shown in Figure 3.E. That plot shows that the differences are not distributed around zero, rather, for some observations of AG2034, the differences are significantly less than zero, in the range of $(-7, -4)$ on the $\log(\text{dose}+\delta)$ scale with $\delta=2.74 \times 10^{-6}$, i.e., in the range of $0.001 \mu\text{M}$ to $0.018 \mu\text{M}$ on the original dose scale. Therefore, the pure additive effect model could not describe the data well. We used bivariate thin plate splines to fit the differences versus the transformed doses, with the knots at all the distinct transformed dose levels. The transformation is taken as $\log(\text{dose}+\delta)$ for all single and combination doses. By convention, there is no drug interaction when a single drug is applied. Therefore the differences were set to zero for the marginal doses. Applying the bivariate thin plate splines (Section 3.3), we obtained $\hat{\sigma}_e^2 = 0.0041$, $\hat{\sigma}_u^2 = 0.2318$, and $\hat{\lambda} = \hat{\sigma}_e^2 / \hat{\sigma}_u^2 = 0.0178$. Next, we constructed 95% point-wise upper and lower confidence surfaces for the fitted bivariate spline function $f(d_1, d_2)$ based on equation ($E9$). Figure 3.F shows the contour plot of the fitted spline function $f(d_1, d_2)$ at the levels of $-0.1, 0$, and 0.1 as thin solid lines; the intercept lines of the corresponding 95% point-wise upper confidence surface with the dose plane as thick dashed lines; and the intercept lines of the corresponding 95% point-wise lower confidence surface with the dose plane as thick solid lines. The combination doses inside the thick dashed curves, shown in light blue, are synergistic because the effects beyond additivity at these combination doses are significantly smaller than zero. The combination doses inside the thick solid curves, shown in light pink, are antagonistic because the effects beyond additivity at these combinations were significantly larger than zero. The combination doses in the uncolored region, which lie between the thick solid curves and the thick dashed curves, are additive because the effects beyond additivity are not significantly different from zero. In particular, the combination doses with AG2034 in the transformed scale in the range of $(-7, -4)$ inside the thick dashed lines are synergistic, which is consistent with the residual plot in Figure 3.E. The fitted response surface was obtained by adding the fitted spline function f (i.e., the effect beyond additivity) to the predicted additive surface. The contour plot of the fitted response surface at the contour levels of $0.2, 0.5$, and 0.9 is shown in Figure 3.I. The final residuals were obtained by subtracting the fitted effects from the observed effects. The plots of final residuals versus the dose levels of TMQ and AG2034 on the $\log(\text{dose}+\delta)$ scale are shown in Figures 3.G and 3.H, respectively. From these two plots, we see that the residuals are centered around zero along the experimental dose range. We conclude that the model fits the data reasonably well.

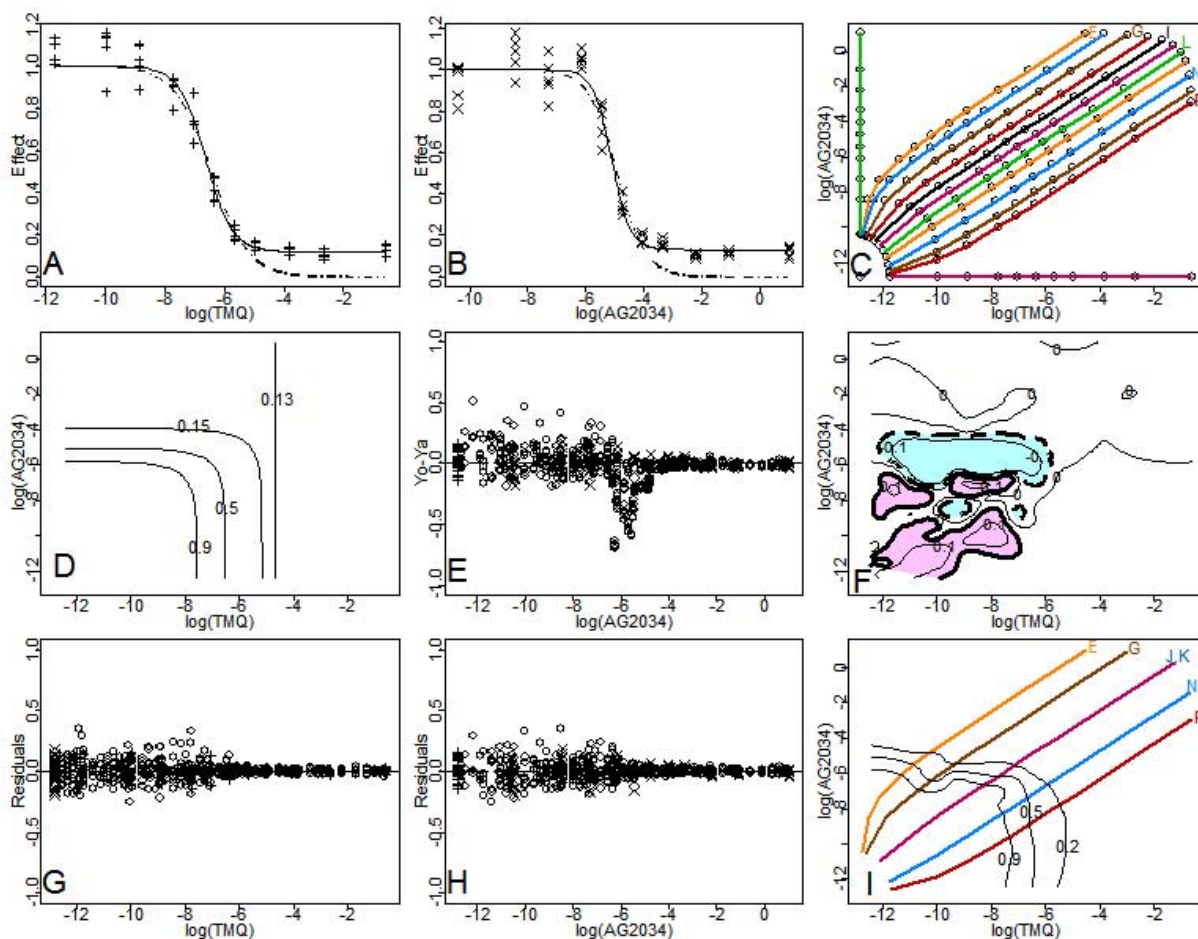


Figure 3. Analysis of the low FA experimental data. Panels A and B show the fitted marginal dose-effect curves for TMQ and AG2034 respectively, where the dotted-dashed line in each panel is the fitted dose-effect curve based on the median effect equation (E 3), while the solid line in each panel is the fitted dose-effect curve based on the E_{max} model (E 5). Panel C shows the distribution of the experimental doses and combination doses on the $\log(\text{dose}+\delta)$ scale with $\delta=2.74 \times 10^{-6}$, along with the 12 rays from left to right with dose ratios of TMQ versus AG2034 at 1:250, 1:125, 1:50, 1:20, 1:10, 1:5, 1:5, 2:5, 4:5, 2:1, 5:1, 10:1, denoted by the letters E, F, G, H, I, J, K, L, M, N, O, and P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, and 14 in the original data set. Panel D shows the contour plot of the predicted additive effect, while Panel E shows the plot of the differences between the observed effects and the predicted effects versus the dose level of AG2034 on the $\log(\text{dose}+\delta)$ scale. Panel F shows the contour plot of the fitted effect beyond the additivity effect at levels -0.1, 0, and 0.1 as thin solid lines, along with the intercept line of the 95% point-wise upper confidence surface with the dose plane as thick dashed lines and the intercept line of the 95% point-wise lower confidence surface with the dose plane as thick solid lines. In Panel F, the combination doses in the light blue area are synergistic, the combination doses in the light pink area are antagonistic, and the combination doses in the uncolored area are additive. The colored lines in Panels C and I are the design rays. Panels G and H are the plots of the final residuals versus TMQ and AG2034 on the $\log(\text{dose}+\delta)$ scale, respectively, and Panel I is the contour plot of the fitted response surface at the levels of 0.9, 0.5, and 0.2, along with some representative design rays.

To examine the patterns of drug interactions in different rays and different experimental combination doses, we combined Figures 3.F and 3.I, that is, we plotted the contour curves of the fitted response surface at the levels of 0.2, 0.5, and 0.9 in Panel F. We also plotted the representative design rays, with the experimental combination doses shown as dots on these rays (Figure 4.A). As seen in that figure, the combination doses on the rays E through K (curves 15, 13, 11, 7, 5, 3, 9 in the original data set) are synergistic when the effect levels are between 0.9 and a number smaller than 0.2. The

combination doses on these rays are additive when the effect level is less than this small number, and the combination doses at low levels on these lines are either additive or antagonistic. The combination doses on the rays N, O, and P (curves 10, 12, and 14 in the original data set) are additive when the effects are less than 0.9, and the combination doses at low dose levels are antagonistic.

In addition to the 95% point-wise confidence interval, we constructed the 95% simultaneous confidence

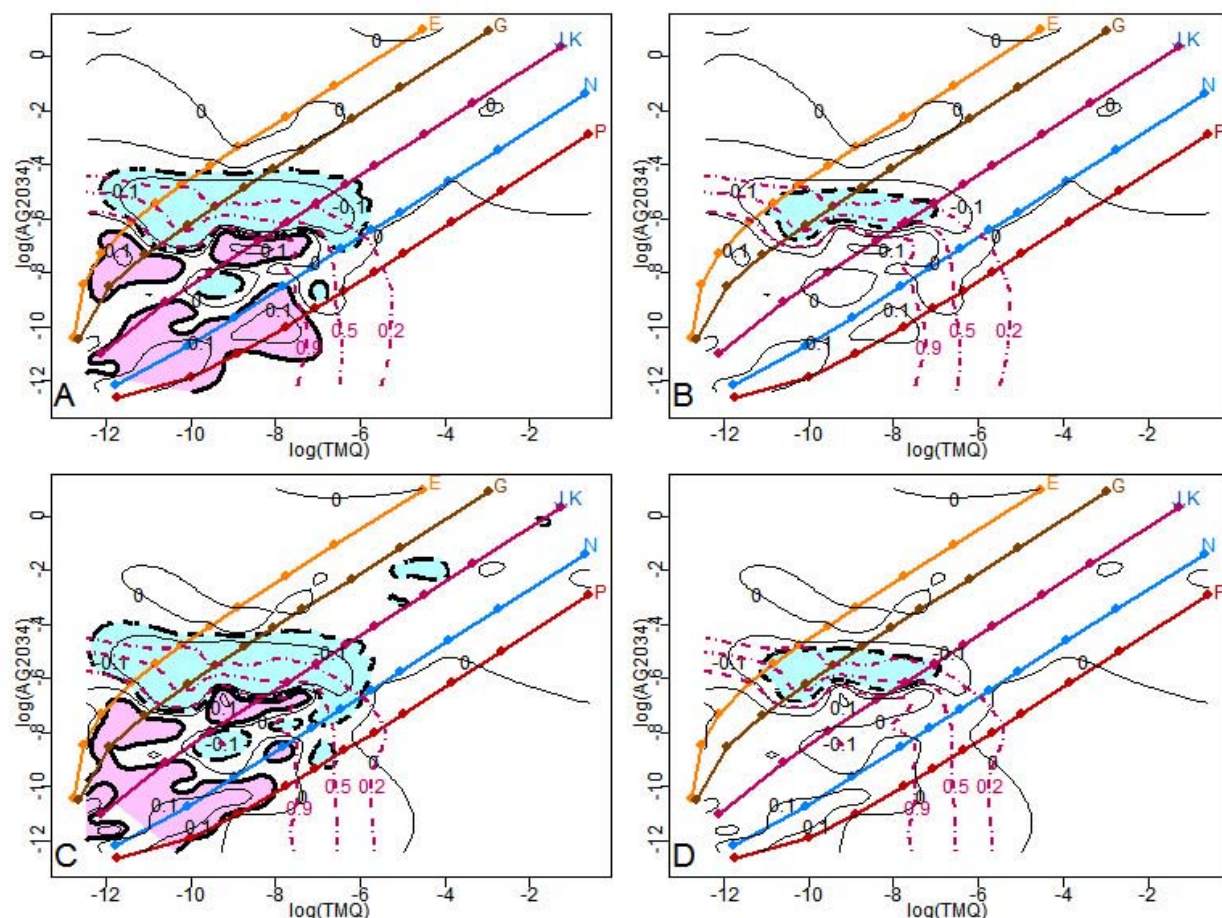


Figure 4. Different patterns of drug interactions for the low FA experimental data. Panel A is based on 95% point-wise confidence intervals; Panel B is based on 95% simultaneous confidence band. Panel A is the combination of Figures 3.F and 3.I, along with the design points shown as dots on each ray. Thin solid lines are contour lines of the fitted effect surface beyond the additivity surface at the levels of -0.1, 0, and 0.1; thick dashed lines are the intercept lines of the upper 95% point-wise confidence surface with the dose plane; thick solid lines are the intercept lines of the lower 95% point-wise confidence surface with the dose plane. The colored lines labeled E, G, J, K, N, and P are the representatives of the design rays. The red dotted-dashed lines are the contour lines of the fitted response surface at the levels of 0.9, 0.5, and 0.2, respectively. Based on Panel A, the combination doses in the light blue area are synergistic, the combination doses in the light pink area are antagonistic, and the combination doses on the uncolored area are additive. Panel B presents the same information as Panel A except that the thick dashed lines are the intercept lines of the upper 95% simultaneous confidence surface with the dose plane and there are no intercept lines for the lower 95% simultaneous confidence surface with the dose plane. Based on Panel B, the combination doses inside the dashed lines are synergistic, otherwise additive. Panel B gives more conservative results for assessing drug interactions. Panels C and D are the results from fitting the data set excluding outliers for the low FA experiment, where the information in Panel C is parallel to that in Panel A, and the information in Panel D is parallel to that in Panel B.

band based on equation (E9) with $z_{\alpha/2}$ replaced by $\sqrt{EDF \times F_{EDF, n-EDF}^{\alpha}}$. Here $EDF=119$, $n=761$, $\sqrt{EDF \times F_{EDF, n-EDF}^{\alpha}} = 12.20$, and $\alpha = 0.05$. The resulting patterns of drug interactions are shown in Figure 4B, in which the thick dashed line is the intercept line of the 95% upper simultaneous confidence surface with the dose plane. Based on Figure 4B, we conclude that the combination doses inside the thick dashed curves, shown in light blue, are synergistic. The combination doses outside the thick dashed curves are additive.

As seen in the figure, compared to the point-wise confidence interval approach, the simultaneous confidence band method shrinks the synergistic area and results in the disappearance of the antagonistic area. A point-wise confidence interval is appropriate for making inferences for each observed combination dose. The simultaneous confidence band is suitable for making a global assessment; however, it can be overly conservative.

Following those analyses, we evaluated the low FA experimental data set from which we had removed the

outliers (19). The results of our assessment of drug interactions for this data set are presented in Figures 4.C and 4.D. Figures 4.A and 4.C contain parallel information, as do Figures 4.B and 4.D. Comparing the plots across the panels, we conclude that the results from fitting the original data set and those from fitting the data set excluding outliers are very similar. Therefore, the semiparametric method presented in Section 3 is robust to outliers in this example.

We recommend using caution when considering extrapolations based on spline estimations. The fitted response surface for the differences between the observed effects and predicted effects gives an overall picture of the drug interactions (see Figures 4.A and 4.B). However, the fitted results on the two larger areas outside experiment rays E and P should not be over-interpreted because (i) there are no experimental data in those areas and (ii) we forced the differences of the observed effects and predicted additive effects to be zero at the marginal observed dose levels.

4.2. Case study 2: cancer cells grown in a medium with 78 μM folic acid (high FA experiment)

In the high FA experiment, the dose levels for TMQ when applied alone were 5.47×10^{-6} , 4.38×10^{-5} , 1.38×10^{-4} , 4.38×10^{-4} , 8.75×10^{-4} , 1.75×10^{-3} , 3.5×10^{-3} , 7×10^{-3} , 2.21×10^{-2} , 7×10^{-2} , and $0.56 \mu\text{M}$, and the dose levels for AG2034 when applied alone were 2.71×10^{-4} , 2.17×10^{-3} , 6.87×10^{-3} , 2.17×10^{-2} , 4.34×10^{-2} , 8.68×10^{-2} , 1.74×10^{-1} , 3.47×10^{-1} , 1.1 , 3.47 , and $27.8 \mu\text{M}$. The procedure we used to analyze this data set was the same as that used in case study 1. By applying nonlinear least squares regression, we estimated the marginal dose-effect curves using the median effect equation (E3) (dotted-dashed lines) and the E_{\max} model (E5) (solid lines), (shown in Figures 5.A and 5.B). It is clear that the E_{\max} model (E5) fits the data better than the median effect equation, thus, we chose the E_{\max} model (E5) as the dose-effect curve for this data set. The estimated parameters for the marginal dose-effect curves for the E_{\max} model are shown in the three columns under the title “High FA” in Table 1. The combination doses on the original scale (not shown) are crowded in the region of the low doses, thus we applied the transformation in the form of $\log(\text{dose}+\delta)$ to each dose level, where δ is a small number, say 2.74×10^{-6} , one half of the lowest dose level for TMQ and AG2034 when applied alone. The distribution of the experimental dose levels on the $\log(\text{dose}+\delta)$ scale is shown in Figure 5.C. The 12 design rays for the combination doses correspond to the 12 dose ratios of TMQ versus AG2034 at 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50, 1:50, 1:25, 1:12.5, 1:5, 1:2, and 1:1, which are denoted by the letters E, F, G, H, I, J, K, L, M, N, O, and P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set for the high FA experiment. Applying the procedure described in Section 3.2, we obtained the contour plot of the predicted additive effect that is shown in Figure 5.D. We see that the contour line at level 0.15 is the predicted effect produced by TMQ alone because AG2034 could not produce such an effect when applied alone; the effect levels for AG2034 applied alone range from 0.1816 to 1. Figure 5.E shows the

differences of the observed effects and predicted effects versus the dose levels of AG2034 on the $\log(\text{dose}+\delta)$ scale. That plot shows that the differences are not centered around zero, rather for some observations of AG2034, the differences are significantly less than zero, in the range of $(-5, 0)$ on the $\log(\text{dose}+\delta)$ scale, i.e., in the range of $6.7 \times 10^{-3} \mu\text{M}$ to $1.0 \mu\text{M}$ on the original dose scale. These findings indicate that some combination doses were synergistic and that the pure additive effect model could not describe the data well. We used bivariate thin plate splines to fit these differences versus the transformed doses or combination dose, with the knots at all the distinct transformed dose levels. The transformation is taken as $\log(\text{dose}+\delta)$ for all single doses and combination doses. We constructed its 95% point-wise confidence surfaces based on equation (E9). The estimated $\hat{\sigma}_\varepsilon^2 = 0.0066$, $\hat{\sigma}_u^2 = 0.0779$, and $\hat{\lambda} = \hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2 = 0.0842$. Figure 5.F shows the contour plot of the fitted spline function f at the levels of -0.1 , 0 , and 0.1 as thin solid lines; the intercept lines of its corresponding 95% point-wise upper confidence surface with the dose plane as thick dashed lines; and the intercept lines of its corresponding 95% point-wise lower confidence surface with the dose plane as thick solid lines. The combination doses inside the thick dashed curves, shown in light blue, are synergistic; the combination doses inside the thick solid curves, shown in light pink, are antagonistic; and the combination doses in the uncolored area are additive. We obtained the fitted response surface by adding the fitted spline function f to the predicted additive surface (shown in Figure 5.I). The plots of the final residuals versus the dose levels of TMQ and AG2034 on the $\log(\text{dose}+\delta)$ scale are shown in Figures 5.G and 5.H, respectively. From these two plots, we see that the residuals are centered around zero along the experimental dose range, indicating that the model describes the data reasonably well.

To examine the patterns of drug interactions in different rays and different experimental combination doses, we combined the plots in Figures 5.F and 5.I to form Figure 6.A (as we did when analyzing the data from the low FA experiment). From Figure 6.A, we see that the combination doses on all 12 rays are synergistic when the effect levels are between 0.9 and 0.15. The combination doses at high dose levels are additive, and most of the combination doses at low dose levels are additive. In addition, we constructed a 95% simultaneous confidence band based on equation (E9) with $z_{\alpha/2}$ replaced by $\sqrt{EDF \times F_{EDF, n-EDF}^\alpha}$. Here $EDF=91$, $n=769$, and $\sqrt{EDF \times F_{EDF, n-EDF}^\alpha} = 10.77$. The results are presented in Figure 6.B, in which the thick dashed line is the intercept line of the upper 95% simultaneous confidence surface with the dose plane. Based on Figure 6.B, we conclude that the combination doses inside the thick dashed curves, shown in light blue, are synergistic. The combination doses outside the thick dashed curves are additive. As in our analysis of the data from case study 1, in this analysis, we found the simultaneous confidence band to yield more conservative results and to be more suitable for a global assessment. For this case study, we also assessed the data set from which we had removed the outliers. The results for assessing drug

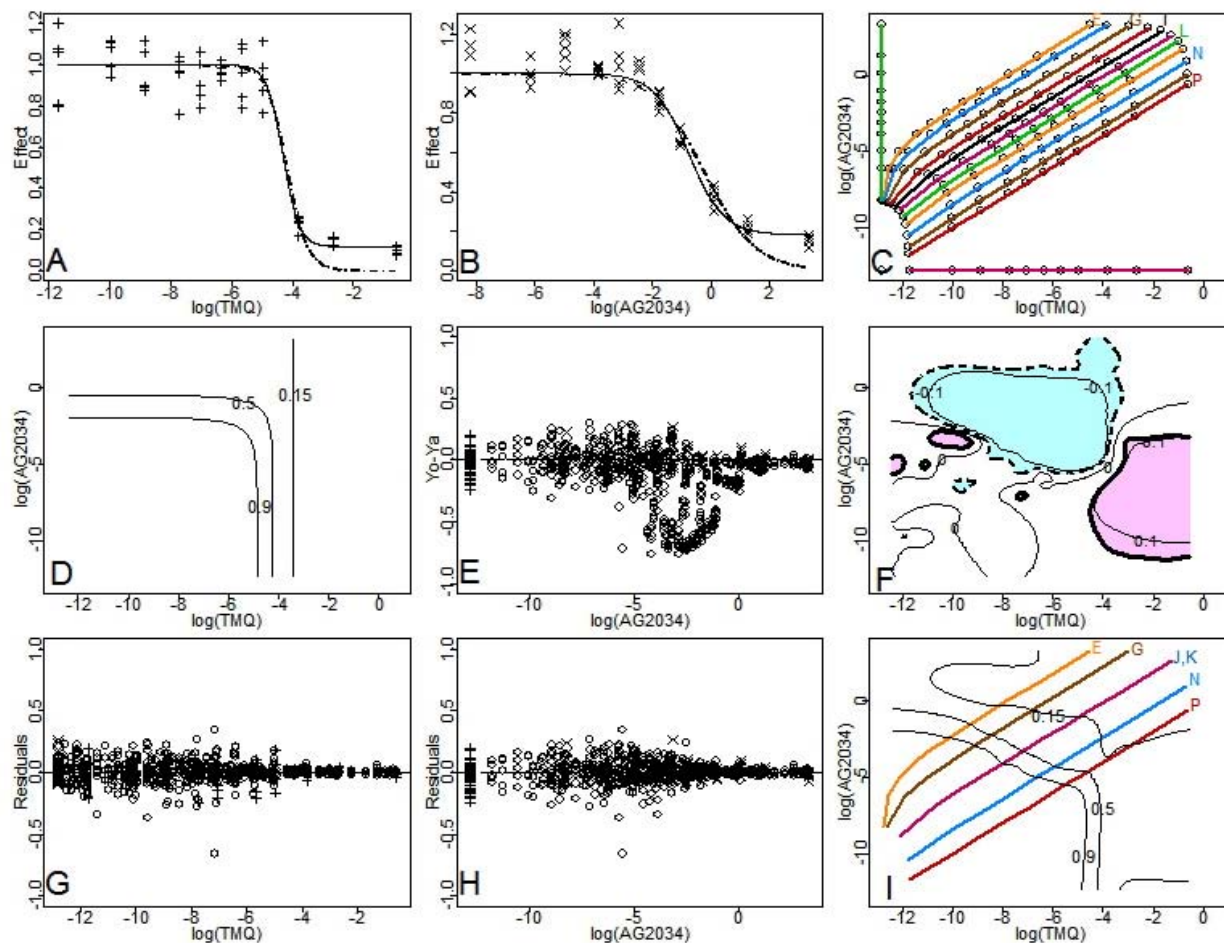


Figure 5. Analysis of the high FA experimental data. Panels A and B show the fitted marginal dose-effect curves for TMQ and AG2034 respectively; the dotted-dashed line is the fitted dose-effect curve based on the median effect equation (E 3); the solid line is the fitted dose-effect curve based on the E_{max} model (E 5). Panel C shows the distribution of the experimental doses and combination doses on the $\log(\text{dose}+\delta)$ scale with $\delta=2.74 \times 10^{-6}$, along with the 12 rays (left to right) with dose ratios of TMQ versus AG2034 at 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50, 1:25, 1:12.5, 1:5, 1:2, 1:1, denoted by the letters E, F, G, H, I, J, K, L, M, N, O, P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set. Panel D shows the contour plot of the predicted additive effect. Panel E shows the plot of the differences between the observed effects and the predicted effects versus the dose level of AG2034 on the $\log(\text{dose}+\delta)$ scale. Panel F shows the contour plot of the fitted effect beyond the additivity effect at levels -0.1, 0, and 0.1, along with the intercept line of the upper 95% point-wise confidence surface with the dose plane as thick dashed lines and the intercept lines of the lower 95% point-wise confidence surface with the dose plane as thick solid lines. In Panel F, synergistic combination doses are in light blue; antagonistic combination doses are in light pink; additive combination doses are in the uncolored area. The colored lines in Panels C and I represent the design rays. Panels G and H are the plots of the final residuals versus TMQ and AG2034 on the $\log(\text{dose}+\delta)$ scale, respectively. Panel I is the contour plot of the fitted response surface at the levels of 0.9, 0.5, and 0.15, along with some representative design rays.

interactions are presented in Figures 6.C and 6.D. Comparing Figures 6.A to C and 6.B to D, we conclude that the results from fitting the original data set and those from fitting the data set excluding outliers are very similar. Thus, the results indicate that the semiparametric method is robust to outliers.

5. SUMMARY AND PERSPECTIVE

We extended the approach proposed by Kong and Lee (6) to a situation for which the E_{max} model is more

appropriate to describe the marginal dose-effect relationship. We considered the possibility that some effect readings at low doses may fall beyond the mean of the controls. Under such circumstances, the standardized effect is greater than 1 and a logit transformation to a linear model (4, 8, 9) cannot be carried out. Hence, other models such as the E_{max} model are needed and nonlinear least squares regression methods can be applied for estimating parameters for the dose-effect curves. We applied nonlinear least squares regression in the case studies to estimate the parameters for the dose-effect curves specified by the median effect equation and the E_{max} model.

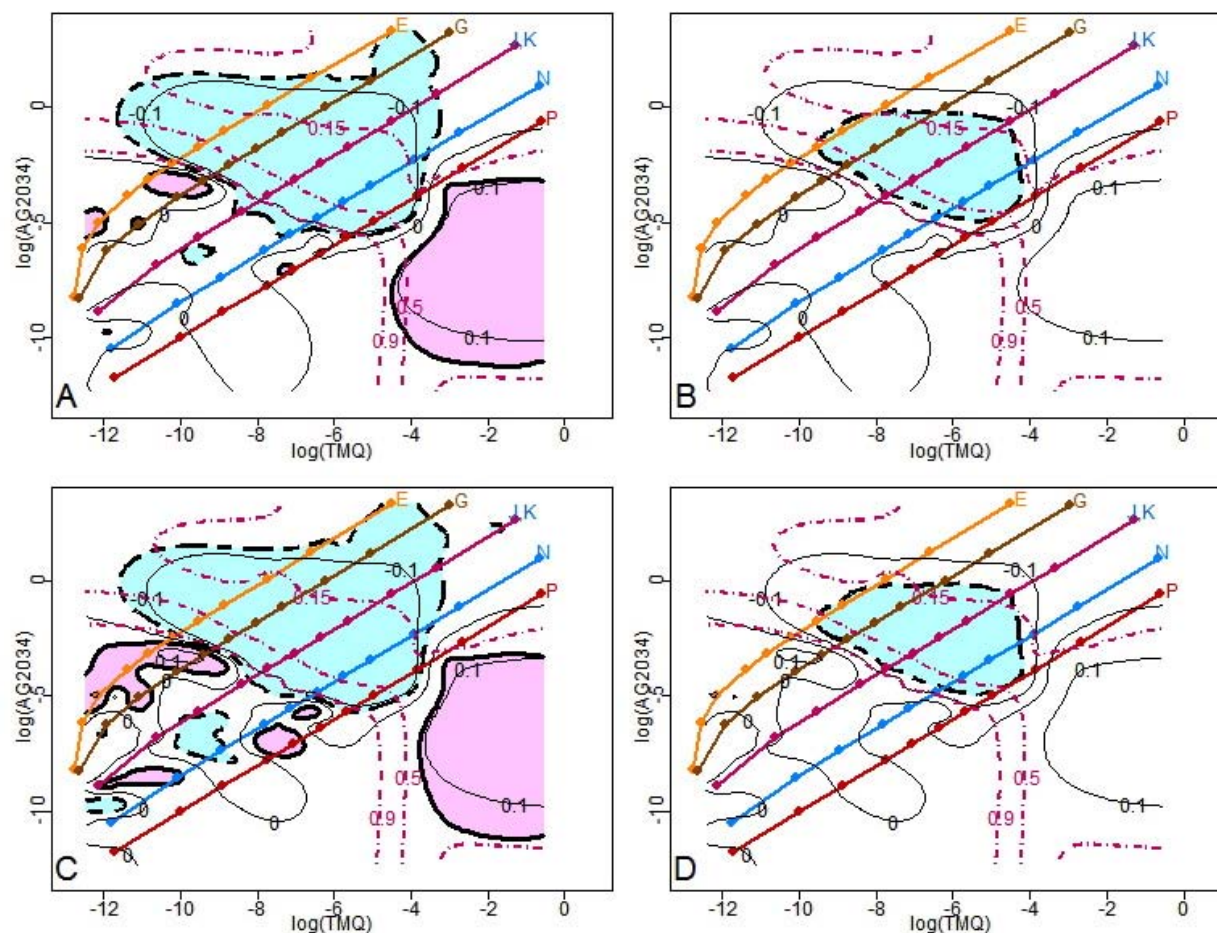


Figure 6. Different patterns of drug interactions for the high FA experiment. Panel A is based on 95% point-wise confidence intervals; Panel B is based on 95% simultaneous confidence band. Panel A is the combination of Figures 5.F and I, along with the design points shown as dots on each ray. Thin solid lines are contour lines of the fitted effect surface beyond the additivity surface at the levels of -0.1, 0, and 0.1; thick dashed lines are the intercept lines of the upper 95% point-wise confidence surface with the dose plane; thick solid lines are the intercept lines of the lower 95% point-wise confidence surface with the dose plane. The colored lines labeled E, G, J, K, N, and P represent the design rays; red dotted-dashed lines are the contour lines of the fitted response surface at the levels of 0.9, 0.5, and 0.15. In Panel A, the synergistic combination doses are in light blue; the antagonistic combination doses are in light pink; additive combination doses in the uncolored area. Panel B gives the same information as Panel A except that the thick dashed lines are the intercept lines of the upper 95% simultaneous confidence surface with the dose plane. Based on Panel B, the combination doses inside the dashed lines are synergistic, otherwise additive. Panel B gives more conservative results for assessing drug interactions. Panels C and D are the results excluding outliers, and are parallel to the results in Panels A and B, respectively.

Additionally, we extended the approach of Kong and Lee (6) as a solution to the problem arising when the experimental points are very close and the low rank of the design matrix may cause computational problems in matrix inversion. For this situation, we considered a low-rank thin plate spline (14) to estimate the surface beyond additivity, and, alternatively, we applied an appropriate transformation to the doses so that the combination doses on the transformed scale were more evenly distributed. In the case studies, we first applied the transformation $\log(\text{dose} + \delta)$ to each component of the combination doses and then applied bivariate thin plate splines with knots representing all the different observed doses on the $\log(\text{dose} + \delta)$ scale. In both

case studies, we chose δ as half of the smallest non-zero dose among TMQ and AG2034 when applied alone, that is, $\delta = 2.74 \times 10^{-6}$ for both experiments. The value of δ selected should not be too small or too large compared with the magnitude of the dose levels. An extremely small δ results in a relatively large distance between the marginal doses and combination doses. Conversely, a large δ dominates in the transformation $\log(\text{dose} + \delta)$ when the dose levels are low. From the final residual plots, it is evident that the current transformation works well.

It is well known that the smoothing parameter λ governs the trade-off between the goodness-of-fit and the smoothness of the function f . When λ becomes larger, the

fitted function f tends to be smoother and the residuals tend to be larger. The selection of the smoothing parameter plays a key role in the fitted results. In the case studies, the smoothing parameter, $\hat{\lambda}$, was selected as $\hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$, which is almost identical to the selected smoothing parameter based on the generalized cross validation criterion and "leave-out-one" cross validation criterion. For example, for the low FA experimental data, the selected parameters based on the mixed model approach used in this paper, cross validation criterion, and generalized cross validation criterion were 0.0178, 0.0112, and 0.0071, respectively. For the high FA experimental data, the corresponding selected parameters were 0.0842, 0.0842, and 0.0531, respectively. Indeed, Kohn, Ansley, and Tharm (20) showed that the estimation of the smoothing parameter based on a mixed model approach is comparable with the standard method of the generalized cross validation criterion. By applying a mixed effects model, the smoothing parameter can be automatically determined by $\hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$. This method has been implemented in S-PLUS by Ruppert *et al* (15) using the lme function (21). In our previous study (1), based on extensive simulations, we showed that the selection of the smoothing parameter provides a good estimate to the underlying function in general.

In the two case studies, we also performed the same analyses for the two reduced data sets analyzed by Lee *et al* (19), and achieved almost identical results, which indicates that this semiparametric method is robust to outliers. The semiparametric method we have developed can also be used to assess drug interactions for the combination doses not on the design rays, and to identify complex patterns of drug interaction in combination studies. In addition, the method gives an overall assessment of the combination effect in the entire two-dimensional dose space spanned by the experimental doses with a caveat that extrapolation beyond data points can be risky.

We also note that the estimated function $f(d_1, d_2)$ and its 95% confidence surfaces can guide the exploration of whether some parametric models are sufficient to describe the data. Many parametric models have been proposed in the literature. Greco, Bravo, and Parsons (3) provided an excellent review of the response surface approach. However, without prior knowledge of the response surface model, or adequate representation of the data by a parametric model, most parametric approaches will fail. Blindly using any parametric model can be dangerous and may lead to wrong conclusions of drug interactions. In our proposed approach, there is no need to assume any parametric models for $f(d_1, d_2)$. We provide a promising approach by modeling the mixture effect data with spline techniques via a mixed-effect model. We advocate the use of the semiparametric method for model building because the true patterns of drug interactions are typically not known. The conclusions regarding drug interactions are based on the estimated f and its confidence surfaces, which are determined by the underlying data. The S-PLUS code used to evaluate the case studies can be obtained from the first author.

6. ACKNOWLEDGEMENT

We thank Dr. William R. Greco at Roswell Park Cancer Institute for supplying the data sets and for his invitation to submit this paper. We thank the two referees for their constructive comments. MK is supported in part by a National Institutes of Health grant P20RR024489. JJJ is supported in part by the Department of Defense grants W81XWH-05-2-0027 and W81XWH-07-1-0306, a National Cancer Institute grant CA16672, and the John G. & Marie Stella Kenedy Foundation Chair in Cancer Research.

7. REFERENCES

1. Fumihiko Kanzawa, Kazuto Nishio, Kazuya Fukuoka, Toshihiko Sunami, and Nagahiro Saijo: *In vitro* interactions of a new derivative of spicamycin, KRN5500, and other anticancer drugs using a three-dimensional model. *Cancer Chemother Pharmacol* 43, 353-363 (1999)
2. Morris C. Berenbaum: What is synergy? *Pharmacol Rev* 41, 93-141 (1989)
3. William R. Greco, Gregory Bravo, John C. Parsons: The search of synergy: A critical review from a response surface perspective. *Pharmacol Rev* 47(2), 331-385 (1995)
4. Ting-Chao Chou: Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol Rev* 58, 621-681 (2006)
5. J. Jack Lee, Maiying Kong, Gregory D. Ayers, Reuben Lotan: Interaction index and different methods for determining drug interaction in Combination Therapy. *J Biopharm Stat* 17, 461-480 (2007)
6. Maiying Kong, J. Jack Lee: A semiparametric response surface model for assessing drug interactions. *Biometrics* 64, 396-405 (2008)
7. Maiying Kong, J. Jack Lee: A general response surface model with varying relative potency for assessing drug interactions. *Biometrics* 62 (4), 986-995 (2006)
8. J. Jack Lee, Maiying Kong: A confidence interval for interaction index for assessing multiple drug interaction. *Stat Biopharm Res* 1, 4-17 (2009)
9. Ting-Chao Chou, Paul Talalay: Quantitative analysis of dose effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv Enzyme Regul* 22, 27-55 (1984)
10. Naitee Ting: Dose Finding in Drug Development. Springer, New York, USA, 127-145 (2006)
11. Daniël M. Jonker, Sandra A. G. Visser, Piet H. van der Graaf, Rob A. Voskuyl, and Meindert Danhof:

Towards a mechanism based analysis of pharmacodynamic drug-drug interaction *in vivo*. *Pharmacol and Ther* 106, 1-18 (2005)

Louisville, Kentucky 40292, U.S.A. Tel: 502-852-3988, Fax: 502-852-3294, E-mail: maiying.kong@louisville.edu

<http://www.bioscience.org/current/vol2E.htm>

12. Peter J. Green, Bernard W. Silverman: Nonparametric Regression and Generalized Linear Models. Chapman & Hall, London (1994)

13. Ming Tan, Hong-Bin Fang, Guo-Liang Tian, and Peter J. Houghton: Experimental design and sample size determination for testing synergism in drug combination studies based on uniform measures. *Stat Med* 22, 2091-2100 (2003)

14. Haonan Wang, and M. Giovanna Ranalli: Low-rank smoothing splines on complicated domains. *Biometrics* 63, 209-217 (2007)

15. David Ruppert, Matthew P. Wand, and Raymond J. Carroll: Semiparametric Regression. Cambridge University Press, UK (2003)

16. Yuedong Wang: Mixed effect smoothing spline analysis of variance. *J Roy Stat Soc B* 60, 159-174 (1998).

17. Henry Scheffe: The Analysis of Variance. New York: John Wiley & Sons (1959)

18. Hélène M. Faessel, Harry K. Slocum, Robert C. Jackson, Theodore J. Boritzki, Youcef M. Rustum, M. G. Nair, and William R. Greco: Super *in vitro* synergy between inhibitors of dihydrofolate reductase and inhibitors of other folate-requiring enzymes: The critical role of polyglutamylolation. *Cancer Res* 58, 3036-3050 (1998)

19. J. Jack Lee, Heather Y. Lin, Diane D. Liu, and Maiying Kong: Applying E_{max} model and interaction index for assessing drug interaction in combination studies. *Front Biosci* (2009) [in press]

20. Robert Kohn, Craig F. Ansley, David Tharm: The performance of cross validation and maximum likelihood estimators of spline smoothing parameters. *J Am Stat Assoc* 86, 1042-1050 (1991)

21. José C. Pinheiro, Douglas M. Bates: Mixed-Effects Models in S and S-PLUS. New York: Springer-Verlag (2000)

Abbreviations: ANOVA = analysis of variance
FA = folic acid
HCT-8 cells = human illeocecal adenocarcinoma cells
lme = linear mixed effects
μM = micromole
TMQ = trimetrexate
SRB = sulforhodamine B

Key Words: Additivity, Antagonism, Bivariate splines, Combination therapy, E_{max} model, the Loewe additivity model, Synergy, Review

Send correspondence to: Maiying Kong, Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville,

E_{max} model and interaction index for assessing drug interaction in combination studies

J. Jack Lee,¹ Heather Y. Lin,¹ Diane D. Liu,¹ Maiying Kong²

¹Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Unit 1411, 1515 Holcombe Boulevard, Houston, Texas 77030, U.S.A., ²Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, Kentucky 40292, U.S.A.

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Statistical methods
 - 3.1. Data sets
 - 3.2. E_{max} model
 - 3.3. Interaction index under the E_{max} model
 - 3.4. Confidence interval for the interaction index
 - 3.5. Data analysis plan
4. Exploratory data analysis
5. Data preprocessing: outlier rejection and data standardization
 - 5.1. Outlier rejection
 - 5.2. Data standardization
6. Results
 - 6.1. Results for the low folic acid experiment
 - 6.2. Results for the high folic acid experiment
7. Summary
8. Discussion and Perspective
9. Acknowledgement
10. References

1. ABSTRACT

Applying the E_{max} model in a Lowe additivity model context, we analyze data from a combination study of trimetrexate (TMQ) and AG2034 (AG) in media of low and high concentrations of folic acid (FA). The E_{max} model provides a sufficient fit to the data. TMQ is more potent than AG in both low and high FA media. At low TMQ:AG ratios, when a smaller amount of the more potent drug (TMQ) is added to a larger amount of the less potent drug (AG), synergy results. When the TMQ:AG ratio reaches 0.4 or larger in low FA medium, or when the TMQ:AG ratio reaches 1 or larger in high FA medium, synergy is weakened and drug interaction becomes additive. In general, synergistic effect in a dilution series is stronger at higher doses that produce stronger effects (closer to $1-E_{max}$) than at lower dose levels that produce weaker effects (closer to 1). The two drugs are more potent in the low compared to the high FA medium. Drug synergy, however, is stronger in the high FA medium.

2. INTRODUCTION

Due to complex disease pathways, combination treatments can be more effective and less toxic than treatments with a single drug regimen. Successful applications of combination therapy have improved treatment effectiveness for many diseases. For example, the combination of a non-nucleoside reverse transcriptase inhibitor or protease inhibitor with two nucleosides is considered a standard front-line therapy in the treatment of AIDS. Typically, a combination of three to four drugs is required to provide a durable response and reconstitution of the immune system (1). Another example is platinum-based doublet chemotherapy regimens as the standard of care for patients with advanced stage non-small-cell lung cancer (2). Combination treatments have also been shown to prevent and to overcome drug resistance in infectious diseases such as malaria and in complex diseases such as cancer (3, 4). Emerging developments in cancer therapy involve combining multiple targeted agents with or without

chemotherapy, or combining multiple treatment modalities such as drugs, surgical procedures, and/or radiation therapy (5, 6).

How do we assess the effect of a combination therapy? It is a simple question, yet it requires a complex answer. A superficial way to answer the question is to determine that a combination therapy is working if its effect is greater than that produced by each single component given alone. The notion of classifying drug interaction as additive, synergistic, or antagonistic is logical and easily understood in a general sense, but can be confusing in specific application without consensus on a standard definition. Excellent reviews of drug synergism have been written by Berenbaum (7), Greco *et al* (8), Suhnel (9), Chou (10), and Tallarida (11), to name a few. In essence, to quantify the effect of combination therapy, we must first define drug synergy in terms of “additivity.” An effect produced by a combination of agents that is more (or less) than the additive effect of the single agents is considered synergistic (or antagonistic). Then, we must further assess drug interaction in a statistical sense. Under a more rigorous definition, synergy occurs when the combined drug effect is statistically significantly higher than the additive effect. Conversely, antagonism occurs when the combination effect is statistically significantly lower than the additive effect.

Despite the controversy arising from multiple definitions of additivity or no drug interaction, the Loewe additivity model is commonly accepted as the gold standard for quantifying drug interaction (7-11). The Loewe additivity model is defined as

$$\frac{d_1}{D_{y,1}} + \frac{d_2}{D_{y,2}} = 1. \quad (E1)$$

Here y is the predicted additive effect at the combination dose (d_1, d_2) when the two drugs do not interact. $D_{y,1}$ and $D_{y,2}$ are the respective doses of drug 1 and drug 2 required to produce the same effect y when used alone. Note that the Loewe additivity can be easily demonstrated in a “sham combination” (i.e., a drug combined with itself or its diluted form). For example, suppose drug 2 is a 50% diluted form of drug 1. The combination of one unit of drug 1 and one unit of drug 2 will produce the same effect as 1.5 units of drug 1 or 3 units of drug 2. Plugging the respective values in equation (E1), we have $1/1.5 + 1/3 = 1$. Given the dose-effect relationship for each single agent, say $E_i(d) = f_i(d)$ for agent i ($i=1,2$), $D_{y,i}$ can be obtained by using the inverse function of f_i , say, $f_i^{-1}(y)$. Replacing $D_{y,1}$ and $D_{y,2}$ in equation (E1) with $f_1^{-1}(y)$ and $f_2^{-1}(y)$, respectively, we can rewrite equation (E1) as

$$\frac{d_1}{f_1^{-1}(y)} + \frac{d_2}{f_2^{-1}(y)} = 1. \quad (E2)$$

Note that (E2) involves an unknown variable y . By solving equation (E2), the predicted additive effect y_{add} can be obtained under the Loewe additivity model. Denote that the observed mean effect is y_{obs} at the combination dose (d_1, d_2). The drug combination at that dose is considered synergistic, additive, or antagonistic when the effect y_{obs} is

greater than, equal to, or less than y_{add} , respectively. When the dose-effect curve is decreasing (or increasing), a synergistic effect corresponds to a smaller (or larger) value than the predicted quantity.

Alternatively, to measure and quantify the magnitude of drug interaction, the interaction index (II) can be defined as

$$II = \frac{d_1}{D_{y_{obs},1}} + \frac{d_2}{D_{y_{obs},2}} \quad (E3)$$

Note that $II < 1$, $II = 1$, and $II > 1$ correspond to the drug interaction being synergistic, additive, and antagonistic, respectively. Chou and Talalay (12) proposed the following median effect equation (E4) to characterize the dose-effect relationship in combination studies:

$$E(d) = \frac{(d / ED_{50})^m}{1 + (d / ED_{50})^m}, \quad (E4)$$

where ED_{50} is the dose required to produce 50% of the maximum effect. Although the median effect equation can be applied in many settings, it assumes that when m is positive, $E(d)=0$ for $d=0$ and $E(d)=1$ for $d=\infty$. On the other hand, when m is negative, $E(d)=1$ for $d=0$ and $E(d)=0$ for $d=\infty$. If we assume that the data follow the median effect equation, a linear relationship can be found by plotting the logit transformation of the effect versus the logarithm transformed dose. A more detailed account of the interpretation and use of the interaction index can be found in a number of references (13-16). Several methods for constructing the confidence interval estimation of the interaction index were proposed by Lee and Kong (17).

To help advance research developing and comparing methods for analyzing data for combination studies, Dr. William R. Greco at the Roswell Park Cancer Institute has organized an effort and invited several groups to participate in an exercise to compare rival modern approaches to model data from two-agent concentration-effect studies. We describe the data and statistical methods, including the E_{max} model, and the calculation of the interaction index under the E_{max} model in Section 3. We describe an exploratory data analysis in Section 4, and data preprocessing for outlier rejection and standardization in Section 5. We present the main results of the data analysis in Section 6 and summarize our findings in Section 7. We close with a discussion in Section 8.

3. STATISTICAL METHODS

3.1. Data sets

Two data sets provided by Dr. Greco are used to examine the effect of the combination treatment of trimetrexate (TMQ) and AG2034 (AG) in HCT-8 human ileocecal adenocarcinoma cells. The cells were grown in a medium with two concentrations of folic acid: 2.3 μ M (the first data set, called low FA) and 78 μ M (the second data set, called high FA). Trimetrexate is a lipophilic inhibitor

of the enzyme dihydrofolate reductase; and AG2034 is an inhibitor of the enzyme glycinamide ribonucleotide formyltransferase. The experiment was conducted on 96-well plates. The endpoint was cell growth measured by an absorbance value (ranging from 0 to 2) and recorded in an automated 96-well plate reader. Each 96-well plate included 8 wells as instrumental blanks (no cells); the remaining 88 wells received drug applications. The experiments were performed using the “ray design,” which maintains a fixed dose ratio between TMQ and AG in a series of 11 dose dilutions. With 88 wells in each plate, each 5-plate stack allowed for an assessment of the combination doses at 7 curves (i.e., design rays) plus a “curve” with all controls. Two stacks were used for studying 14 design rays: TMQ only, AG only, and twelve other design rays with a fixed dose ratio (TMQ:AG) for each ray. The fixed dose ratios in the low FA experiment were 1:250, 1:125, 1:50, 1:20, 1:10, 1:5 (2 sets), 2:5, 4:5, 2:1, 5:1, and 10:1. The fixed dose ratios in the high FA experiment were 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50 (2 sets), 1:25, 2:25, 1:5, 1:2, and 1:1. Data from each of the 16 curves (2 for controls, 2 for single agents, and 12 for combinations) are grouped together. Curves 1-8 were performed on the first stack with curve 8 serving as the “control” experiment while curves 9-16 were performed on the second stack with curve 16 serving as the “control” experiment. The assignment of different drug combinations to the cells in the wells was randomized across the plates. Five replicate plates were used for each set of two stacks, resulting in a total of 10 plates for each of the two medium conditions (low FA and high FA). The maximum number of treated wells per medium condition is 880 (16 curves x 11 dilutions x 5 replicates). Complete experimental details and mechanistic implications were reported by Faessel *et al* (18).

3.2. E_{max} model

Due to a plateau of the measure of cell growth such that it does not reach zero at the maximum dose levels used in the experiments, the median effect equation (E4) does not fit the data. Instead, we take the E_{max} model (19) to fit the data at hand.

$$E(d) = E_0 - E_{max} + \frac{E_{max}}{1 + (d / ED_{50})^m}, \quad (E5)$$

where E_0 is the base effect, corresponding to the measurement of cell growth when no drug is applied; E_{max} is the maximum effect attributable to the drug; ED_{50} is the dose level producing half of E_{max} ; d is the dose level that produces the effect $E(d)$, and m is a slope factor (Hill coefficient) that measures the sensitivity of the effect within a dose range of the drug. Thus, $E_0 - E_{max}$ is the asymptotic effect when a very large dose of the drug is applied. Figure 1 shows a few examples of the E_{max} model where E_0 is assumed to be 1. The parameter m governs how quickly the curve drops. For the three cases in the first row in Figure 1, ED_{50} is fixed at 2 and E_{max} is at 0.8, while the slope varies. When $m=1$ (Figure 1.A), the dose-response

curve drops slowly; when $m=5$ (Figure 1.B and E), a sigmoidal curve is formed, and when $m=20$ (Figure 1.C and F), the drop of the sigmoidal curve becomes very steep. In the three curves in the first row, as the dose increases, the curves drop, and the effect asymptotes to $1 - E_{max} = 0.2$. In the second row, the three plots are set at $E_{max} = 1$, which means that as the dose increases, the treatment will reach the theoretical full effect. For example, if the measure of the treatment effect is cell count, all the cells will be killed at very high doses of the treatment when $E_{max} = 1$. The figures also show that, as ED_{50} increases, the curves shift to the right, indicating that the treatment is less potent. In all cases when m increases, the effect drops more rapidly. We apply the nonlinear weighted least squares method to estimate the parameters in the E_{max} model. Due to the heteroscedascity observed in the data, which means that the variance increases as the observed response increases, we use the reciprocal of the fitted response as the weight function (20). We use S-PLUS, R (21), and SAS (22) to carry out the estimation.

3.3. Interaction index under the E_{max} model

As when using the median effect model, the E_{max} model can be applied to fit the single-drug and combination drug dose-response curves, and then the interaction index can be calculated accordingly. Although equation (E5) allows for different values of E_0 and E_{max} for different curves, when calculating the interaction index, we need to assume all curves have the same E_0 so that the “base measure” of no drug effect is the same in all curves. This can be achieved by dividing all of the effect measures with the mean of the controls. Note that E_{max} can vary in different curves to signify different drug potencies. However, the calculation of the interaction index will be a little more complicated when different drugs or combinations produce different values of E_{max} .

Hereafter, we assume the dose-response curve follows the E_{max} model given in (E6):

$$E(d) = 1 - E_{max} + \frac{E_{max}}{1 + (d / ED_{50})^m}. \quad (E6)$$

The experiments we analyzed studied the ability of the combination treatments to inhibit the growth of cancer cells. The measure of the treatment effect was cell growth corresponding to the number of cells observed. Hence, the height of the dose-effect curve decreases when the dose increases. In this case, we have $m > 0$. In addition, as d goes to infinity, the effect plateaus at $1 - E_{max}$. Hence, E_{max} must be between 0 and 1.

In a study of two-drug combinations, we need to fit three curves using the E_{max} model: curve 1 for drug 1 alone, curve 2 for drug 2 alone, and curve c for the drug combinations. Denote $E_{max, i}$, $ED_{50, i}$, and m_i as the three parameters for drug i ($i=1,2, c$). Given an effect e ($e > 1 - E_{max}$), the corresponding dose $d(e)$ can be calculated as

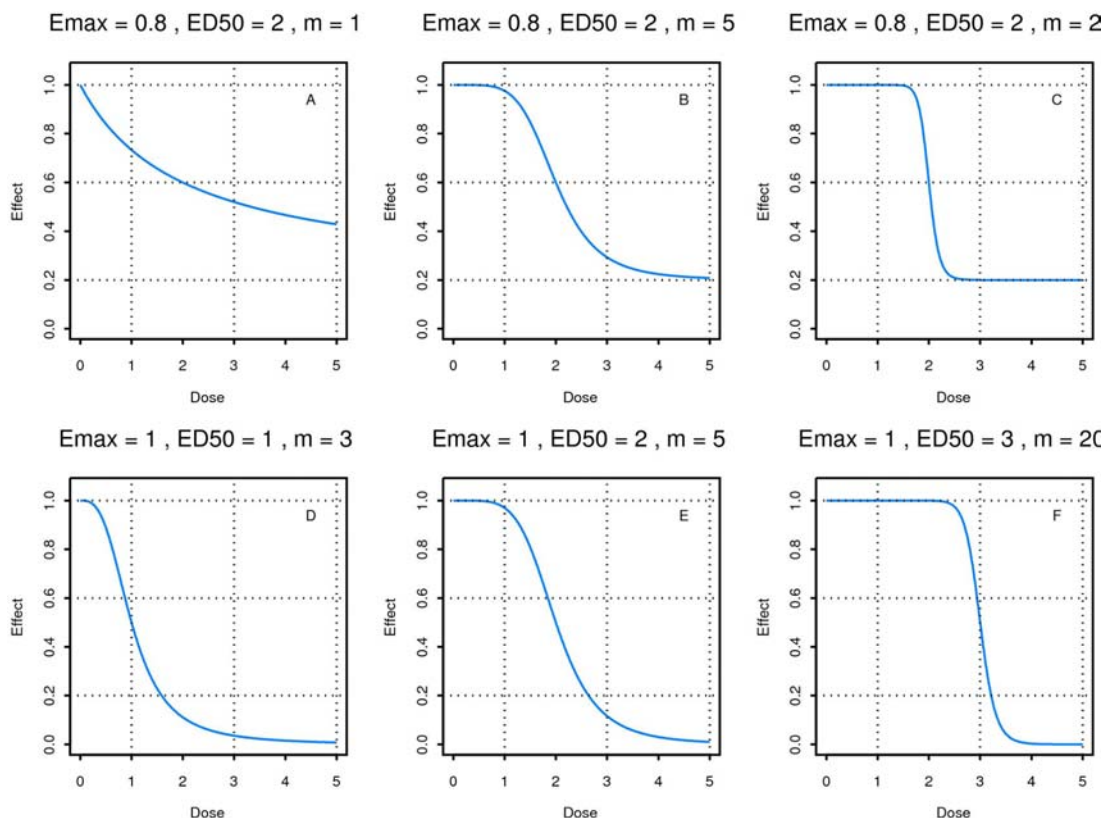


Figure 1. Dose-response curves under the E_{max} model by varying the parameters E_{max} , ED_{50} , and m .

$$d(e) = ED_{50} \left(\frac{1-e}{e-1+E_{max}} \right)^{1/m} \quad (E7)$$

Note that the dose for the combination treatment is simply the sum of the doses of the single agents. This approach works well for a ray design with constant or varying relative potency between the two drugs (12, 17). Without loss of generality, we can assume that $E_{max,1} > E_{max,2}$. In addition, we assume that the dose ratio for the two drugs in the combination treatment ($d_c = d_1 + d_2$) is fixed with $d_1/d_2 = p$. Upon fitting the three dose-response curves, the interaction index at a fixed effect e where $e \in (1 - \hat{E}_{max,c})$ be calculated as follows:

$$\hat{II} = \frac{\hat{d}_c(e) \times p / (1+p)}{\hat{D}_{y,1}(e)} + \frac{\hat{d}_c(e) / (1+p)}{\hat{D}_{y,2}(e)} \text{ for } 1 - \hat{E}_{max,2} < e < 1, \text{ and} \quad (E8)$$

$$\hat{II} = \frac{\hat{d}_c(e) \times p / (1+p)}{\hat{D}_{y,1}(e)} \text{ for } 1 - \hat{E}_{max,1} < e \leq 1 - \hat{E}_{max,2}.$$

For $e \leq 1 - \hat{E}_{max,1}$, the interaction index cannot be calculated. However, the combination effect in this range is more than additive because it reaches an effect level that no single agent can achieve alone. If $E_{max,1} = E_{max,2}$, the interaction index can be calculated using the first formula in (E8).

3.4. Confidence interval for the interaction index

We can apply the delta method to calculate the (large sample) variance of the interaction index (23). From our previous work (17), we found that better estimation of

the confidence interval for the interaction index can be achieved by working on the logarithmic transformation of the interaction index.

$$\text{By applying the delta method, } \text{Var}(\log(\hat{II})) = \frac{1}{\hat{II}^2} \text{Var}(\hat{II}).$$

When $1 - \hat{E}_{max,2} < e < 1$, the variance of \hat{II} can be calculated by

$$\text{Var}(\hat{II}) = \left(\frac{\hat{d}_c(e) \times p / (1+p)}{\hat{D}_{y,1}(e)} \right)^2 V_1 + \left(\frac{\hat{d}_c(e) / (1+p)}{\hat{D}_{y,2}(e)} \right)^2 V_2 + \left(\frac{\hat{d}_c(e) \times p / (1+p)}{\hat{D}_{y,1}(e)} + \frac{\hat{d}_c(e) / (1+p)}{\hat{D}_{y,2}(e)} \right)^2 V_c \quad (E9)$$

if $1 - \hat{E}_{max,1} < e \leq 1 - \hat{E}_{max,2}$, then

$$\text{Var}(\hat{II}) = \left(\frac{\hat{d}_c(e) \times p / (1+p)}{\hat{D}_{y,1}(e)} \right)^2 (V_1 + V_c). \quad (E10)$$

where

$$V_i = \left(\frac{-1}{\hat{m}_i(e-1+\hat{E}_{max,i})} - \frac{1}{ED_{50,i}} - \frac{1}{\hat{m}_i^2} \log \frac{1-e}{e-1+\hat{E}_{max,i}} \right) \text{Var} \left(\hat{E}_{max,i}, ED_{50,i}, \hat{m}_i \right) \begin{pmatrix} \frac{-1}{\hat{m}_i(e-1+\hat{E}_{max,i})} \\ \frac{1}{ED_{50,i}} \\ \frac{1}{\hat{m}_i^2} \log \frac{1-e}{e-1+\hat{E}_{max,i}} \end{pmatrix}$$

for $i=1, 2, c$.

Upon calculation of the variance for $\log(\hat{II})$, the point-wise $(1-\alpha)100\%$ confidence interval for the interaction index (II) for a specified effect can be constructed as

$$\left(\hat{II} \exp(-z_{\alpha/2} \sqrt{\text{Var}(\log(\hat{II}))}), \hat{II} \exp(z_{\alpha/2} \sqrt{\text{Var}(\log(\hat{II}))}) \right) \quad (E11)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ upper percentile of the standard normal distribution. We also construct the simultaneous confidence band for the interaction index over the range of estimated responses. Because the estimation process involves estimating nine parameters from three curves, to construct a Scheffe-type simultaneous confidence band, we simply replace $z_{\alpha/2}$ in equation (E11) by $(\chi^2_{p-9}(\alpha))^{1/2}$ where $p=9$ (24).

3.5. Data analysis plan

The overall objective of the data analysis is to assess the synergistic effect of the combination of TMQ and AG in both low and high FA media. We apply the exploratory data analysis first, and then estimate the dose-response relationship using the E_{max} model. We evaluate the drug interaction by calculating the interaction index under the Loewe additivity model. We perform an exploratory data analysis in order to understand the data structure and patterns and to determine whether preprocessing of the data in terms of outlier rejection and standardization would be required prior to data modeling. We analyze the low FA and high FA experiments separately then compare the results. For each experiment, we apply the E_{max} model to fit the two marginal and twelve combination dose-response curves. We compute the interaction index and its 95% confidence intervals for each of the twelve combinations, and assess the overall pattern of drug interaction by examining the interaction index from the 12 fixed-ratio combinations together. We apply a one-dimensional distribution plot via the BLiP plot (25) to display the data. We use a two-dimensional scatter plot, a contour plot, and an image plot as well as a three-dimensional perspective plot to show the dose-response relationship. We also apply a trellis plot (26) to assemble the individual plots together into consecutive panels conditioning on different values of fixed dose ratios.

4. EXPLORATORY DATA ANALYSIS

As in all data analyses, we begin with an exploratory data analysis. For the low and high FA experiments, there are 871 and 879 readings, respectively. Only 9 and 1 observations, respectively, are missing out of a maximum of 880 readings in each experiment. The data include designated curve numbers ranging from 1 to 16 and data point numbers ranging from 1 to 176. Each curve number indicates a specific dose combination. We re-label the curves as A-P where A and B correspond to the control (no drug) curves; C and D correspond to the curves of TMQ and AG administered alone, and curves E through P correspond to the combination curves with fixed dose ratios in ascending order. Each point number indicates the readings at each specific dilution of each curve. Because five duplicated experiments were performed, there are up to five readings for each specific point number. There is, however, no designation of the plate number in the data received. Figure 2 shows the variable percentile plot of the distribution of the effect from the low FA and high FA experiments using the BLiP plot, with each segment corresponding to a five percent increment (25). The plot gives an overall

assessment of the distribution of the outcome variable of cell growth without conditioning on experimental settings. The middle 20% of the data (40th to 60th percentiles) are shaded in a light orange color. This figure indicates that the data have a bimodal distribution with most data clustered around either a low value of 0.2 or a high value of 1.2. For the low FA experiment, the distribution of the effect ranges from 0.072 to 1.506 with the lower, middle, and upper quartiles being 0.149, 0.449, and 1.150, respectively. Similarly, for the high FA experiment, the effect ranges between 0.070 and 1.545. The three respective quartiles are 0.213, 0.990, and 1.1495. The median of the data from the low FA experiment is smaller than the median of the data from the high FA experiment. The bimodal distributions could result from steep dose-response curves. As a consequence, the slope may not be estimated well in certain cases.

To help understand the pattern of the fixed ratio dose assignment in a ray design and the relationship between the fixed ratio doses and curve numbers, we plot the logarithm transformed dose of TMQ and AG in Figure 3 for both the low FA and high FA experiments. As can be seen, curves A and B are the controls with no drugs. Curves C and D correspond to the single drug study of TMQ and AG, respectively. Curves E through P are the various fixed ratio combination doses of TMQ and AG. Note that curves J and K have the same dose ratios. Within each curve, the 11 dilutions are marked by 11 circles. For the combination studies, the curves for different dose ratios are parallel to each other on the log dose scale. If the same plot is shown in the original scale, these lines will form “rays,” radiating out from the origin like sun rays. Hence, the term “ray design” is an appropriate name for this type of experiment. The corresponding dose ranges used for each drug alone are 5.47×10^{-6} to $0.56 \mu\text{M}$ for TMQ in both the low FA and high FA experiments, and 2.71×10^{-5} to $2.78 \mu\text{M}$ for AG2034 in the low FA experiment and 2.71×10^{-4} to $27.78 \mu\text{M}$ in the high FA experiment.

Figures 4 and 5 show the raw data of the effect versus dose level by curve for the low FA and high FA experiments, respectively. Instead of using the actual dose, we plot the data using a sequentially assigned dose level to indicate each dilution within each curve such that the data can be shown clearly. In addition, the data points at each dilution for each curve are coded from 1 to 5 according to the order of the appearance in the data set. We assume that these numbers correspond to the replicate number for each design point (the well position in the stack of 5 plates). Because the plate number was not listed in the data, we are not certain that this is the case. From the plot, we can see that there are outliers in several dilution series. Of note, in Figure 4, the effects from plate (replicate) #1 in curves B, E, F, and K tend to be lower than all other replicates. There are also some unusually large values, for example, in replicate 2 in curve A, dose level (dilution series) 6; replicate 3 in curve L, dose level 4; and replicate 2 in curve M, dose level 1. Similar observations can be made for the

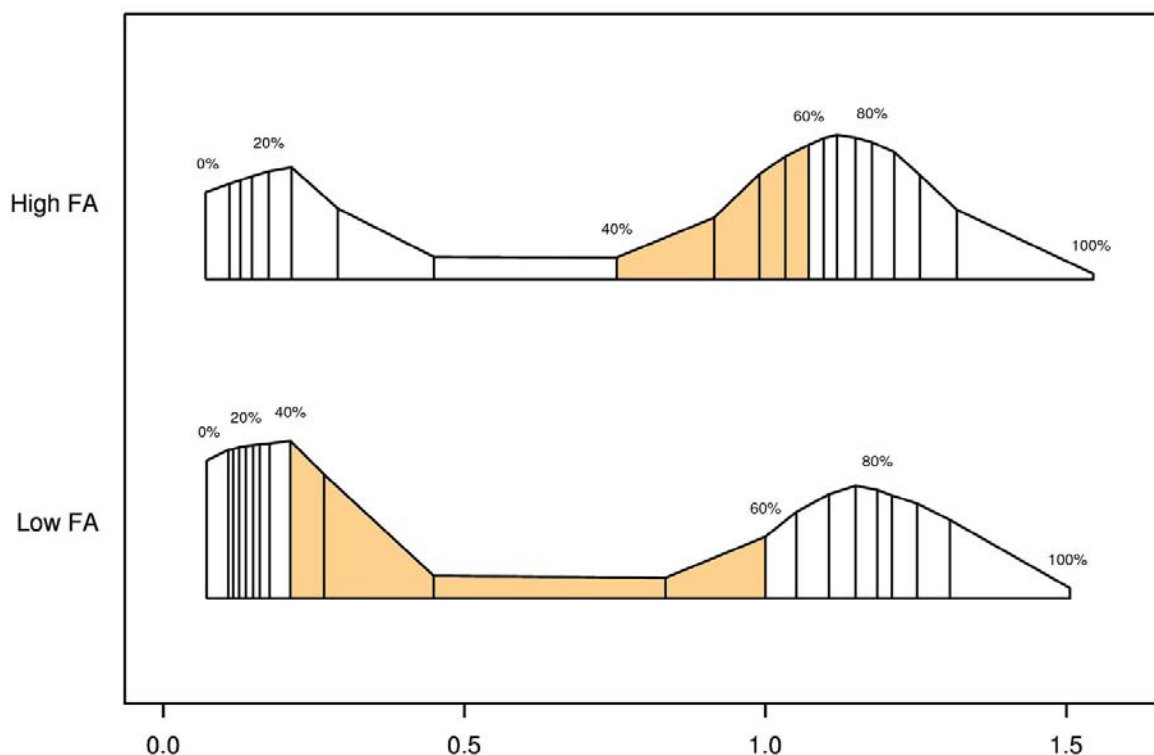


Figure 2. Variable width percentile plot for the observed effect in experiments with low and high folic acid media. Each vertical bar indicates a five percent increment. The middle 20% of the data are shaded in a light orange color.

high FA experiment: plate #1 seems to have some low values in curves B, C, H, I, and J, and plate #4 seems to have some low values in curves E, K, N, O, and P. These findings indicate that certain procedures need to be performed to remove the obvious outliers in order to improve the data quality before the data analysis.

Figure 6 shows the perspective plot, contour plot, and image plot for the low FA experiment. From the perspective plots in Figure 6.A (back view), B (front view), and C (side view), we can see that the effect starts at a high plane plateau at an effect level of about 1.2 when the doses of TMQ are AG are small. As the dose of each drug increases, the effect remains approximately constant for a while and then a sudden drop occurs. This steep downward slope can be found by taking the trajectory of any combination of the TMQ and AG doses; it is also evident in the dose-response curves shown in Figures 4 and 5. The steep drop of the effect can also be found in the contour plot and the image plot. Similar patterns in the dose-response relationship are shown in Figure 7 for the high FA experiment. The steep drop of the effect occurs at smaller doses in the low FA experiment and at larger doses in the high FA experiment.

5. DATA PREPROCESSING: OUTLIER REJECTION AND DATA STANDARDIZATION

5.1. Outlier rejection

To address the concern that outliers may adversely affect the analysis outcome, we devise the following simple

plan. For each of the 176 point numbers (16 curves x 11 dilutions), the five effect readings should be close to each other because they are from replicated experiments. However, because the plate number is not in the data set, we cannot assess the plate effect. Neither can we reject a certain replicate plate entirely should there be a plate with outlying data, nor apply a mixed effect model treating the plate effect as a random effect. For the four or five effect readings in each point number (only 9 point numbers in the low FA and 1 in the high FA experiments have 4 readings), we compute the median and the interquartile range. An effect reading is considered an outlier if the value is beyond the median ± 1.4529 times the interquartile range. If the data are normally distributed (i.e., follow a Gaussian distribution), the range expands to cover the middle 95% of the data. Hence, only about 5% of the data points (2.5% at each extreme) are considered outliers. The number 1.4529 is obtained by $qnorm(.975)/(qnorm(.75) - qnorm(.25))$ where $qnorm(x)$ is a quantile function which returns the x^{th} percentiles from a normal distribution. Upon applying the above rule, 129 out of 871 (14.8%) effect readings in the low FA experiment and 126 out of 879 (14.3%) in the high FA experiment are considered outliers and are removed before proceeding to further analysis. The numbers of outliers in replicates 1 to 5 are 60, 28, 19, 14, and 8 for the low FA experiment and 35, 18, 21, 34, and 18 for the high FA experiment, indicating a non-random pattern of outliers that could be attributed to experimental conditions. Note that the outlier rejection algorithm is only applied “locally.” In other words, it only applies to the replicated readings up to five replicates in each of the 176 experimental conditions.

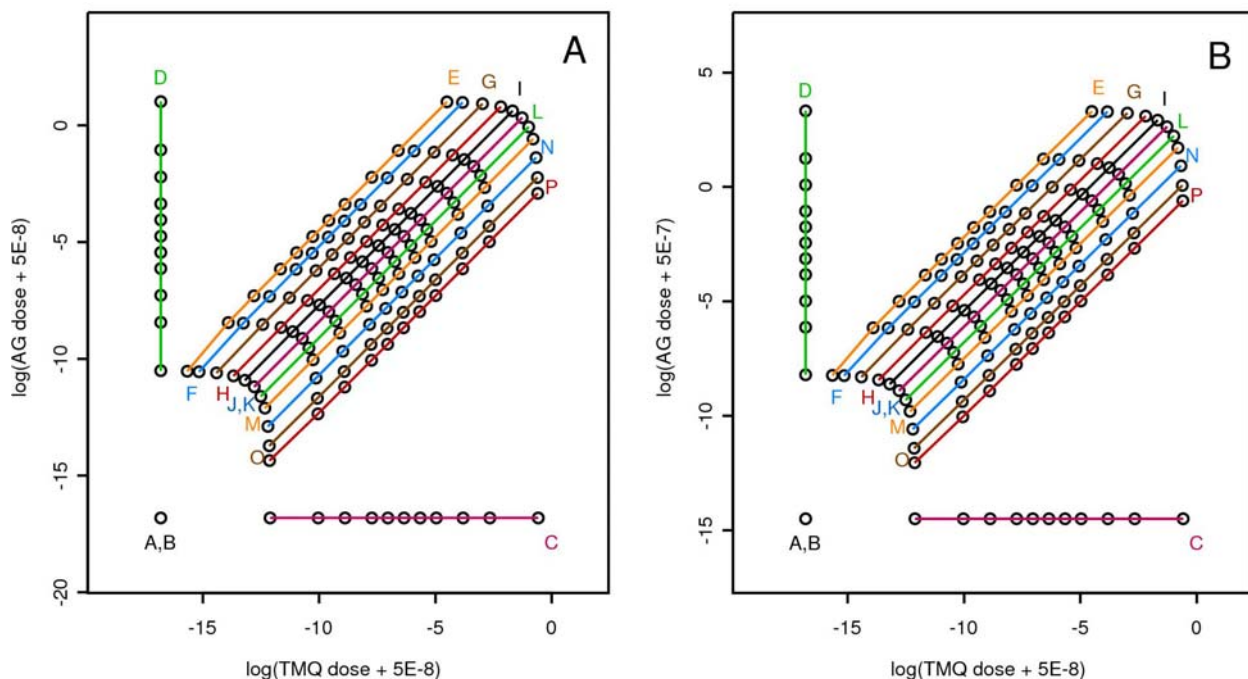


Figure 3. Experimental design showing the logarithmically transformed AG2034 (AG) dose versus the logarithmically transformed trimetrexate (TMQ) dose in the fixed ratio experiments. 16 curves are shown. Curves A and B are controls; no drugs applied. Curves C and D are single-drug studies for TMQ and AG, respectively. Curves E through P are the combination drug studies. Each curve has 11 dilutions shown in circles. Panel A: low folic acid medium. Panel B: high folic acid medium.

5.2. Data standardization

After outliers are removed from the data, we compute the mean of the control curves. The means for curves 8 and 16 are 1.1668 and 1.1534 for the low FA experiment and 1.1483 and 1.1477 for the high FA experiment, respectively. To apply the E_{max} model in equation (E6) with $E_0 = 1$, we standardize the data by dividing the effect readings of respective curves 1-7 by the mean of curve 8 and the effect readings of respective curves 9-15 by the mean of curve 16.

6. RESULTS

6.1. Results for the low folic acid experiment

The E_{max} model in equation (E6) is applied to fit all of the dose-response curves. For the low FA experiment, the parameter estimates, their corresponding standard errors, and the residual sum of squares are given in Table 1. The dose-response relationships showing the data and the fitted curves are displayed in Figure 8. Note that although model fitting is performed on the original dose scale, the dose is plotted on the logarithmically transformed scale to better show the dose-response relationship. The fitted marginal dose-response curves for TMQ (curve C) and AG (curve D) are shown in a blue dashed line and a red dotted line, respectively. From Table 1, we see that \hat{ED}_{50} is 0.00133 for TMQ and 0.00621 for AG, indicating that TMQ is about 4.7 times more potent than AG at the ED_{50} level. For

curves E through P, the fitted dose-response curve for the combination treatment is shown as a solid black line superimposed on the marginal dose-response curves. The proposed E_{max} model fits all curves well except for curves G, H and K. For curve G, although the model estimates converge in an initial attempt, the parameter m is estimated with a standard error of 30.3. The large standard error essentially indicates that the estimate \hat{m} is not reliable. For curve K, the model does not converge on the original dose scale but converges on the logarithmically transformed dose scale. However, the standard error of the estimate \hat{m} is still very large, which leads us to believe that the model is not very stable. For curve H, as can be seen in Figure 8, there are no observed effects between 0.3 and 1 from the second to the fifth dilutions. The parameter m cannot be estimated and the model fails to converge on both the original scale and the logarithmic scale. To address these problems, we conclude that the data do not provide us sufficient information to yield a reasonable estimate of the parameter m . Therefore, we take a remedial approach by fixing m , and then proceed to estimate the other two parameters. Upon checking the data, we set the parameter m as 5, 4.5, and 5 for curves G, H, and K, respectively. The choice of m is somewhat arbitrary with a goal of yielding a good fit to the data and producing a small residual sum of squares. The resulting “reduced” models fit the data reasonably well but with a consequence that there is no standard error estimate for \hat{m} , which affects the variance estimation of the interaction index (to be shown later).

Table 1. Summary of parameter estimates (standard error) for the low FA experiment

Curve ¹	Dose ratio (TMQ/AG) ²	E _{max} ³	ED ₅₀ ⁴	m ⁵	Residual sum of squares
C (1)		0.877 (0.007)	0.00133 (0.00006)	2.345 (0.190)	0.0779
D (2)		0.872 (0.007)	0.00621 (0.00024)	3.045 (0.269)	0.0749
E (15)	0.004	0.869 (0.008)	0.00359 (0.00017)	3.250 (0.437)	0.0969
F (13)	0.008	0.863 (0.008)	0.00294 (0.00014)	2.621 (0.276)	0.0897
G (11*)	0.02	0.865 (0.006)	0.00151 (0.00005)	5.0	0.0817
H (7*)	0.05	0.889 (0.007)	0.00274 (0.00011)	4.5	0.1025
I (5)	0.1	0.885 (0.005)	0.00253 (0.00009)	3.449 (0.306)	0.0689
J (3)	0.2	0.882 (0.005)	0.00244 (0.00007)	4.019 (0.402)	0.0655
K (9*)	0.2	0.872 (0.007)	0.00233 (0.00007)	5.0	0.0843
L (4)	0.4	0.889 (0.006)	0.00278 (0.00011)	5.473 (0.583)	0.0855
M (6)	0.8	0.890 (0.005)	0.00200 (0.00007)	3.208 (0.263)	0.0738
N (10)	2	0.887 (0.008)	0.00169 (0.00009)	2.544 (0.258)	0.0984
O (12)	5	0.878 (0.008)	0.00145 (0.00007)	2.206 (0.206)	0.0837
P (14)	10	0.874 (0.006)	0.00134 (0.00006)	1.971 (0.128)	0.0599

Footnotes and abbreviations: ¹curves without data: curves A(8) and B(16) represent controls; no drugs applied, ²trimetrexate/experimental drug AG2034, ³maximum effect attributable to the drug, ⁴dose level producing half of E_{max}, ⁵slope factor, which measures the sensitivity of the effect within a dose range of the drug, is fixed at a certain value

Based on limited sensitivity analysis, the estimation of the interaction index remains reasonably robust.

In all dose-response curves, the standardized effect level starts to drop between dose levels (dilutions) 3 to 6. Once the effect starts to drop, it drops quickly and plateaus at the $1 - \hat{E}_{max}$ level. There are ample data points at the effect levels around 1 (dose levels 1-4) and $1 - \hat{E}_{max}$ (dose levels 8-11). However, due to the sharp drop in the dose-response curves, fewer data points can be found in the middle of the effect range. When the data points become too few or do not spread out to cover enough range, it becomes harder for the model to converge, as seen in curves G, H, and K. The overall results for the curve fitting of the low FA experiment are that the values of \hat{E}_{max} range

from 0.863 to 0.890; the values of \hat{ED}_{50} range from 0.00133 to 0.00621; and the values of \hat{m} range from 1.971 to 5.473. The residual sum of squares ranges from 0.0599 to 0.1025 without large values, suggesting that the model fits the data reasonably well.

Based on the fitted dose-response curve, the interaction index (II) can be calculated over the entire effect range and at specific dose combinations. Table 2 gives a detailed result of the estimated interaction index and its 95% point-wise confidence interval at each dose combination for each combination curve. The II is calculated at the predicted effect level from the combination curve and not at the observed effect level. The results are shown in a trellis plot in Figure 9 where the red lines represent the 95% point-wise confidence intervals at each specific effect level and the black dashed lines indicate the 95% simultaneous confidence bands of the II for the entire range. From the figure we find that the interaction index can be estimated with very good precision in all curves except at the two extremes when the effect is close to 1 or $1 - \hat{E}_{max}$. The trend and the pattern of the interaction index are clearly shown in these figures. For curves E through K, i.e., with a TMQ:AG dose ratio ranging from 0.004 to 0.2, synergy is observed in the effect range between 0.2 to 0.9. For curves L and M, which have

TMQ:AG ratios of 0.4 and 0.8, we see that synergy is observed at the low effect level from 0.2 to about 0.5. Beyond 0.5 the combinations are generally additive. For curves N, O, and P, with TMQ:AG ratios of 2, 5, and 10, the synergistic effect is lost and we see additivity in all dose ranges.

6.2. Results for the high folic acid experiment

Table 3 lists the parameter estimate, corresponding standard error, and sum of squares for all the curves in the high FA experiment. Unlike in the experiment using low FA media, the model fitting for all curves in the high FA experiment converge when using the E_{max} model.

The estimated \hat{E}_{max} ranges from 0.831 to 0.893; \hat{ED}_{50} ranges from 0.0137 to 0.1943 except for curve D (AG alone with $\hat{ED}_{50} = 0.5224$); and \hat{m} ranges between 1.468 and 3.625. The residuals sum of squares ranges from 0.0615 to 0.1134. Compared to the low FA experiment, the values of \hat{ED}_{50} are greater in the high FA experiment, indicating that the drugs are less potent when applied to a high FA medium. Note that the doses of TMQ are the same between the two experiments but the doses of AG are 10 times

higher in the high FA experiment. In addition, $\hat{ED}_{50} = 0.0137$ and 0.00133 for TMQ alone in the high and low FA experiments, respectively, which indicates that the drug is 10 times less potent in the high FA medium compared to the low FA medium. The potency of AG is even more dramatically reduced. In Figure 10 we see that the E_{max} model provides an excellent fit to all the curves. Table 4 gives a detailed account of the interaction index in all dilutions for all of the combination curves. The results are summarized in a trellis plot in Figure 11. Again, the red lines represent the 95% point-wise confidence intervals at each specific effect level and the black dashed lines correspond to the simultaneous confidence bands of the II for the whole range. Using the high FA medium, synergy can be achieved for most of the drug combinations in all the effect ranges, with the exception of the very low or very high effect ranges. The confidence intervals are still very tight although they are a little wider compared to their counterparts from the low FA experiment. As the TMQ:AG

Emax model and interaction index for drug interaction

Table 2. Estimated interaction index and its 95% confidence interval at each dose combination for the low FA experiment

Curve ¹	TMQ ² dose	AG2034 ³ dose	Dose ratio TMQ/AG	Dilution	Predicted effect	II ⁴	95% CI ⁵ for II	
							Lower limit	Upper limit
E (15)	1.07E-07	2.66E-05	0.004	1	1	0.87	0.18	4.29
	8.58E-07	0.000213		2	0.9999	0.73	0.28	1.89
	2.71E-06	0.000673		3	0.9962	0.67	0.37	1.20
	8.58E-06	0.002129		4	0.864	0.61	0.48	0.78
	1.72E-05	0.004259		5	0.4454	0.58	0.52	0.65
	3.43E-05	0.008517		6	0.1802	0.56	0.45	0.71
	6.86E-05	0.017000		7	0.1368	0.61	0.24	1.55
	0.000137	0.034100		8	0.1319	0.91	0	4.35E+03
	0.000434	0.107700		9	0.1314	2.71	0	1.30E+157
	0.001373	0.340700		10	0.1314	8.58	0	NA
	0.011000	2.725500		11	0.1314	68.6	0	NA
F (13)	2.10E-07	2.61E-05	0.008	1	1	0.28	0.08	0.91
	1.68E-06	0.000209		2	0.9991	0.35	0.17	0.71
	5.32E-06	0.000660		3	0.9828	0.4	0.26	0.62
	1.68E-05	0.002088		4	0.746	0.47	0.4	0.55
	3.37E-05	0.004177		5	0.3788	0.52	0.46	0.58
	6.73E-05	0.008353		6	0.188	0.59	0.47	0.74
	0.000135	0.016700		7	0.1454	0.76	0.4	1.47
	0.000269	0.033400		8	0.138	1.26	0.02	67.11
	0.000851	0.105700		9	0.1366	3.79	0	1.72E+37
	0.002692	0.334100		10	0.1366	11.95	0	NA
	0.021500	2.673100		11	0.1366	95.63	0	NA
G (11)	4.97E-07	2.47E-05	0.02	1	1.0000	5.60	1.90	16.45
	3.98E-06	0.000197		2	1.0000	1.09	0.60	1.96
	1.26E-05	0.000624		3	0.9885	0.47	0.36	0.63
	3.98E-05	0.001974		4	0.2987	0.22	0.19	0.25
	7.95E-05	0.003949		5	0.1410	0.17	0.09	0.32
	0.000159	0.007898		6	0.1350	0.27	0.00	3.09E+04
	0.000318	0.015800		7	0.1348	0.54	0.00	1.62E+161
	0.000636	0.031600		8	0.1348	1.09	0.00	NA
	0.002012	0.099900		9	0.1348	3.44	0.00	NA
	0.006364	0.315900		10	0.1348	10.87	0.00	NA
	0.050900	2.527300		11	0.1348	86.95	0.00	NA
H (7)	1.09E-06	2.17E-05	0.05	1	1.0000	12.11	3.53	41.51
	8.75E-06	0.000174		2	1.0000	2.71	1.42	5.15
	2.77E-05	0.000549		3	0.9992	1.29	0.88	1.90
	8.75E-05	0.001738		4	0.8773	0.65	0.57	0.74
	0.000175	0.003475		5	0.3035	0.43	0.38	0.48
	>=0.000350	>=0.006950		6 - 11	<= 0.1219	NA	NA	NA
I (5)	1.82E-06	1.81E-05	0.1	1	1.0000	2.39	0.70	8.17
	1.46E-05	0.000145		2	0.9999	1.17	0.59	2.30
	4.61E-05	0.000458		3	0.9966	0.83	0.55	1.24
	0.000146	0.001448		4	0.8509	0.61	0.52	0.71
	0.000292	0.002896		5	0.3906	0.51	0.47	0.55
	0.000583	0.005792		6	0.1506	0.38	0.31	0.47
	>=0.001167	>=0.011600		7 - 11	<= 0.1188	NA	NA	NA
	2.73E-06	1.36E-05		1	1.0000	12.56	2.59	60.87
J (3)			0.2					

Emax model and interaction index for drug interaction

	2.19E-05	0.000109		2	1.0000	3.31	1.38	7.95
	6.92E-05	0.000343		3	0.9993	1.67	0.99	2.81
	0.000219	0.001086		4	0.9344	0.88	0.71	1.08
	0.000438	0.002172		5	0.5008	0.61	0.57	0.65
	0.000875	0.004344		6	0.1577	0.40	0.33	0.48
	>=0.001750	>=0.008688		7 - 11	<= 0.1204	NA	NA	NA
K (9)	2.73E-06	1.36E-05	0.2	1	1.0000	88.77	14.80	532.45
	2.19E-05	0.000109		2	1.0000	9.70	3.84	24.48
	6.92E-05	0.000343		3	0.9998	3.04	1.83	5.06
	0.000219	0.001086		4	0.9550	1.02	0.86	1.21
	0.000438	0.002172		5	0.4457	0.55	0.51	0.60
	0.000875	0.004344		6	0.1429	0.32	0.21	0.47
	>=0.001750	>=0.008688		7 - 11	<= 0.1280	NA	NA	NA
L (4)	3.65E-06	9.05E-06	0.4	1	1.0000	812.88	82.77	7.98E+03
	2.92E-05	7.24E-05		2	1.0000	53.71	13.70	210.53
	9.22E-05	0.000229		3	1.0000	12.38	5.18	29.56
	0.000292	0.000724		4	0.9964	2.99	1.98	4.51
	0.000583	0.001448		5	0.8651	1.31	1.10	1.56
	0.001167	0.002896		6	0.2103	0.56	0.50	0.63
	>=0.002333	>=0.005792		7 - 11	<= 0.1134	NA	NA	NA
M (6)	4.38E-06	5.43E-06	0.8	1	1.0000	4.95	1.18	20.69
	3.50E-05	4.34E-05		2	1.0000	2.40	1.01	5.71
	0.000111	0.000137		3	0.9989	1.63	0.93	2.87
	0.000350	0.000434		4	0.9580	1.12	0.85	1.48
	0.000700	0.000869		5	0.7206	0.90	0.80	1.03
	0.001400	0.001738		6	0.2804	0.72	0.64	0.80
	0.002800	0.003475		7	0.1325	0.41	0.25	0.67
	>=0.005600	>=0.006950		8 - 11	<= 0.1128	NA	NA	NA
N (10)	4.97E-06	2.47E-06	2	1	1.0000	1.37	0.32	5.84
	3.98E-05	1.97E-05		2	0.9998	1.17	0.47	2.89
	0.000126	0.000062		3	0.9967	1.08	0.59	1.98
	0.000398	0.000197		4	0.9417	0.99	0.72	1.37
	0.000795	0.000395		5	0.7418	0.95	0.80	1.13
	0.001591	0.000790		6	0.3742	0.90	0.79	1.02
	0.003182	0.001580		7	0.1721	0.81	0.64	1.03
	>=0.006364	>=0.003159		8 - 11	<= 0.1236	NA	NA	NA
O (12)	5.26E-06	1.04E-06	5	1	1.0000	0.67	0.18	2.52
	4.21E-05	8.35E-06		2	0.9995	0.76	0.33	1.76
	0.000133	2.64E-05		3	0.9934	0.82	0.46	1.45
	0.000421	8.35E-05		4	0.9227	0.88	0.65	1.21
	0.000841	0.000167		5	0.7294	0.92	0.78	1.10
	0.001683	0.000334		6	0.4094	0.97	0.87	1.08
	0.003365	0.000668		7	0.2060	1.01	0.83	1.24
	0.006731	0.001337		8	0.1420	1.05	0.67	1.65
	>=0.021300	>=0.004227		9 - 11	<= 0.1239	NA	NA	NA
P (14)	5.36E-06	5.32E-07	10	1	1.0000	0.39	0.14	1.09
	4.29E-05	4.26E-06		2	0.9988	0.54	0.28	1.05
	0.000136	1.35E-05		3	0.9887	0.65	0.42	1.02
	0.000429	4.26E-05		4	0.9015	0.79	0.62	1.01

Emax model and interaction index for drug interaction

0.000858	8.52E-05	5	0.7095	0.88	0.76	1.02
0.001716	0.000170	6	0.4221	0.99	0.89	1.10
0.003431	0.000341	7	0.2272	1.12	0.95	1.32
0.006863	0.000681	8	0.1544	1.30	0.98	1.74
≥ 0.021700	≥ 0.002155	9 - 11	≤ 0.1292	NA	NA	NA

Footnotes and abbreviations: ¹ curves without data: curves A (8) and B (16) represent controls; no drugs applied; curves C (1, TMQ) and D (2, AG2034) represent single-drug applications, ² trimetrexate, ³ experimental drug, ⁴ interaction index, ⁵ confidence interval

Table 3. Summary of parameter estimates (standard error) for the high FA experiment

Curve ¹	Dose ratio (TMQ/AG) ²	E _{max} ³	ED ₅₀ ⁴	m ⁵	Residual sum of squares
C (1)		0.883 (0.012)	0.0137 (0.0012)	3.625 (0.650)	0.1074
D (2)		0.831 (0.015)	0.5224 (0.0439)	1.468 (0.137)	0.0770
E (15)	0.0004	0.867 (0.014)	0.1943 (0.0122)	2.558 (0.405)	0.1134
F (13)	0.0008	0.863 (0.010)	0.1447 (0.0068)	2.643 (0.258)	0.0852
G (11)	0.002	0.859 (0.010)	0.0912 (0.0045)	2.996 (0.355)	0.0999
H (7)	0.005	0.881 (0.006)	0.0699 (0.0027)	2.887 (0.253)	0.0746
I (5)	0.01	0.881 (0.009)	0.0484 (0.0026)	2.528 (0.251)	0.0977
J (3)	0.02	0.884 (0.006)	0.0331 (0.0011)	2.114 (0.136)	0.0615
K (9)	0.02	0.885 (0.008)	0.0369 (0.0019)	2.160 (0.195)	0.0861
L (4)	0.04	0.886 (0.008)	0.0288 (0.0014)	2.504 (0.255)	0.0959
M (6)	0.08	0.885 (0.009)	0.0197 (0.0010)	2.242 (0.214)	0.0881
N (10)	0.2	0.862 (0.010)	0.0154 (0.0007)	3.309 (0.415)	0.0909
O (12)	0.5	0.878 (0.009)	0.0139 (0.0006)	3.491 (0.405)	0.0933
P (14)	1	0.893 (0.008)	0.0183 (0.0009)	2.735 (0.213)	0.0669

Footnotes and abbreviations: ¹ curves without data: curves A(8) and B(16) represent controls; no drugs applied, ² trimetrexate/experimental drug AG2034, ³ maximum effect attributable to the drug, ⁴ dose level producing half of E_{max}, ⁵ slope factor, which measures the sensitivity of the effect within a dose range of the drug

Table 4. Estimated interaction index and its 95% confidence interval at each dose combination for the high FA experiment

Curve ¹	TMQ ² dose	AG2034 ³ dose	Dose ratio TMQ/AG	Dilution	Predicted effect	II ⁴	95% CI ⁵ for II	
							Lower limit	Upper limit
E (15)	1.07E-07	0.000266	0.0004	1	1.0000	48.28	2.53	922.71
	8.58E-07	0.002128		2	1.0000	10.31	1.36	78.32
	2.71E-06	0.006729		3	0.9998	4.39	0.96	20.02
	8.58E-06	0.021278		4	0.9970	1.87	0.68	5.14
	1.72E-05	0.042555		5	0.9825	1.12	0.55	2.28
	3.43E-05	0.085110		6	0.9063	0.67	0.44	1.02
	6.86E-05	0.170221		7	0.6388	0.40	0.32	0.48
	0.000137	0.340441		8	0.2994	0.21	0.16	0.28
	≥ 0.000434	≥ 1.076570		9 - 11	≤ 0.1433	NA	NA	NA
F (13)	2.10E-07	0.000261	0.0008	1	1.0000	42.38	3.74	479.71
	1.68E-06	0.002087		2	1.0000	8.02	1.56	41.25
	5.32E-06	0.006599		3	0.9998	3.20	0.96	10.63
	1.68E-05	0.020868		4	0.9949	1.27	0.59	2.75
	3.37E-05	0.041737		5	0.9688	0.73	0.44	1.23
	6.73E-05	0.083474		6	0.8363	0.42	0.32	0.56
	0.000135	0.166947		7	0.4876	0.24	0.20	0.28
	0.000269	0.333894		8	0.2224	0.11	0.08	0.16
	≥ 0.000851	≥ 1.055866		9 - 11	≤ 0.1418	NA	NA	NA
G (11)	4.97E-07	0.000247	0.002	1	1.0000	80.07	5.84	1097.54
	3.98E-06	0.001973		2	1.0000	9.20	1.67	50.66
	1.26E-05	0.006239		3	0.9997	2.78	0.84	9.26
	3.98E-05	0.019730		4	0.9913	0.85	0.42	1.71
	7.95E-05	0.039460		5	0.9351	0.42	0.28	0.63
	0.000159	0.078920		6	0.6609	0.21	0.17	0.25

Emax model and interaction index for drug interaction

	0.000318	0.157841		7	0.2796	0.10	0.08	0.12
	>=0.000636	>=0.315682		8 - 11	<= 0.1614	NA	NA	NA
H (7)	1.09E-06	0.000217	0.005	1	1.0000	33.80	3.34	342.20
	8.75E-06	0.001736		2	1.0000	4.54	1.03	20.08
	2.77E-05	0.005491		3	0.9994	1.50	0.54	4.20
	8.75E-05	0.017363		4	0.9843	0.51	0.29	0.90
	0.000175	0.034725		5	0.8955	0.27	0.20	0.37
	0.000350	0.069450		6	0.5603	0.15	0.13	0.17
	0.000700	0.138900		7	0.2239	0.07	0.06	0.09
	>=0.001400	>=0.277800		8 - 11	<= 0.1344	NA	NA	NA
I (5)	1.82E-06	0.000181	0.01	1	1.0000	4.94	0.61	39.76
	1.46E-05	0.001447		2	0.9999	1.11	0.30	4.13
	4.61E-05	0.004575		3	0.9977	0.50	0.21	1.20
	0.000146	0.014469		4	0.9592	0.23	0.15	0.37
	0.000292	0.028938		5	0.8071	0.16	0.12	0.20
	0.000583	0.057875		6	0.4556	0.11	0.09	0.13
	0.001167	0.115750		7	0.2041	0.07	0.06	0.09
	>=0.002333	>=0.231500		8 - 11	<= 0.1347	NA	NA	NA
J (3)	2.73E-06	0.000136	0.02	1	1.0000	0.67	0.13	3.34
	2.19E-05	0.001085		2	0.9993	0.28	0.10	0.74
	6.92E-05	0.003432		3	0.9924	0.18	0.09	0.34
	0.000219	0.010852		4	0.9206	0.13	0.09	0.17
	0.000438	0.021703		5	0.7354	0.11	0.09	0.13
	0.000875	0.043406		6	0.4261	0.10	0.09	0.12
	0.001750	0.086813		7	0.2139	0.10	0.08	0.11
	>=0.003500	>=0.173625		8 - 11	<= 0.1405	NA	NA	NA
K (9)	2.73E-06	0.000136	0.02	1	1.0000	0.93	0.15	5.72
	2.19E-05	0.001085		2	0.9995	0.36	0.12	1.12
	6.92E-05	0.003432		3	0.9946	0.22	0.11	0.47
	0.000219	0.010852		4	0.9390	0.15	0.10	0.22
	0.000438	0.021703		5	0.7800	0.13	0.10	0.16
	0.000875	0.043406		6	0.4722	0.11	0.10	0.13
	0.001750	0.086813		7	0.2316	0.11	0.09	0.13
	>=0.003500	>=0.173625		8 - 11	<= 0.1443	NA	NA	NA
L (4)	3.65E-06	0.000090	0.04	1	1.0000	2.89	0.34	24.37
	2.92E-05	0.000723		2	0.9999	0.69	0.18	2.62
	9.22E-05	0.002288		3	0.9983	0.33	0.14	0.80
	0.000292	0.007234		4	0.9702	0.18	0.12	0.29
	0.000583	0.014469		5	0.8538	0.15	0.11	0.19
	0.001167	0.028938		6	0.5312	0.13	0.11	0.15
	0.002333	0.057875		7	0.2326	0.12	0.10	0.15
	>=0.004667	>=0.115750		8 - 11	<= 0.1355	NA	NA	NA
M (6)	4.38E-06	5.43E-05	0.08	1	1.0000	0.73	0.10	5.17
	3.50E-05	0.000434		2	0.9998	0.27	0.08	0.89
	0.000111	0.001373		3	0.9973	0.17	0.08	0.37
	0.000350	0.004341		4	0.9660	0.13	0.09	0.21
	0.000700	0.008681		5	0.8594	0.13	0.10	0.17
	0.001400	0.017363		6	0.5823	0.14	0.12	0.16
	0.002800	0.034725		7	0.2842	0.16	0.13	0.18
	>=0.005600	>=0.069450		8 - 11	<= 0.1571	NA	NA	NA

Emax model and interaction index for drug interaction

N (10)	4.97E-06	2.47E-05	0.2	1	1.0000	61.54	3.00	1262.99
	3.98E-05	0.000197		2	1.0000	4.66	0.64	34.02
	0.000126	0.000624		3	1.0000	1.21	0.31	4.73
	0.000398	0.001973		4	0.9983	0.41	0.20	0.87
	0.000795	0.003946		5	0.9830	0.28	0.17	0.46
	0.001591	0.007892		6	0.8570	0.23	0.17	0.31
	0.003182	0.015784		7	0.4280	0.21	0.18	0.25
	0.006364	0.031568		8	0.1800	0.23	0.18	0.30
	>=0.020124	>=0.099827		9 - 11	<= 0.1390	NA	NA	NA
O (12)	5.26E-06	1.04E-05	0.5	1	1.0000	194.98	6.90	5509.68
	4.21E-05	8.35E-05		2	1.0000	11.38	1.16	111.41
	0.000133	0.000264		3	1.0000	2.57	0.51	12.90
	0.000421	0.000835		4	0.9998	0.78	0.30	1.97
	0.000841	0.001669		5	0.9978	0.50	0.25	0.98
	0.001683	0.003339		6	0.9754	0.40	0.24	0.64
	0.003365	0.006678		7	0.7849	0.36	0.27	0.48
	0.006731	0.013356		8	0.3109	0.35	0.30	0.41
	>=0.021285	>=0.042235		9 - 11	<= 0.1263	NA	NA	NA
P (14)	5.36E-06	5.32E-06	1	1	1.0000	10.38	0.65	165.28
	4.29E-05	4.26E-05		2	1.0000	1.89	0.29	12.19
	0.000136	0.000135		3	1.0000	0.87	0.24	3.15
	0.000429	0.000426		4	0.9998	0.55	0.23	1.29
	0.000858	0.000851		5	0.9986	0.50	0.25	1.03
	0.001716	0.001702		6	0.9910	0.51	0.29	0.92
	0.003431	0.003404		7	0.9436	0.56	0.37	0.86
	0.006863	0.006809		8	0.7232	0.64	0.50	0.83
	0.021702	0.021531		9	0.1851	0.80	0.65	0.99
	>=0.068627	>=0.068088		10, 11	<= 0.1109	NA	NA	NA

Footnotes and abbreviations: ¹curves without data: curves A(8) and B(16) represent controls; no drugs applied; curves C(1, TMQ) and D (2, AG2034) represent single-drug applications, ²trimetrexate, ³experimental drug, ⁴interaction index, ⁵confidence interval

ratio increases from 0.0004 to 0.5, synergy is observed across all dilution series. In addition, higher synergy is observed at the lower effect levels, particularly when the TMQ:AG is at 0.01 or lower (curves E, F, G, H, and I). In the middle effect levels (effects between 0.2 and 0.8), the II ranges from about 0.1 in curves J and K, to 0.12 in curve L, 0.15 in curve M, 0.25 in curve N, and 0.35 in curve O. The higher the TMQ:AG ratio, the less synergy it achieves. In curve P, for example, when the TMQ:AG ratio reaches 1, synergy is lost.

7. SUMMARY

In both the low FA and high FA experiments, TMQ is more potent than AG. At low TMQ:AG ratios, i.e., when a small amount of the more potent drug (TMQ) is added to a larger amount of the less potent drug (AG), synergy is achieved. However, when the TMQ:AG ratio reaches 0.4 or larger for the low FA medium, or when the TMQ:AG ratio reaches 1 or larger for the high FA medium, synergy decreases, or the interaction becomes additive. In general, a synergistic effect in a drug combination dilution series is stronger at higher doses that produce stronger effects (effects closer to $1-E_{max}$) than at lower dose levels

that produce weaker effects (effects closer to 1). The two drugs in this study are more potent in the low FA medium compared to the high FA medium. The drug synergy, however, is stronger in the high FA medium.

8. DISCUSSION AND PERSPECTIVE

The data supplied by Dr. Greco provide an excellent opportunity to apply and compare various approaches for studying the effects of combination drug treatments. For the median effect model, a linear relationship between the logit transformed effect and the log-dose makes the model fitting straightforward and easy. However, when measuring cell growth, as in the experiments we analyzed, if the maximum drug effect reaches a plateau and does not kill all the cancer cells, even at the highest experimental doses, the median effect model (12) does not apply. We used the E_{max} model (19), which provides an adequate fit for most data. Parameter estimation under the E_{max} model requires the use of iterative procedures such as the nonlinear weighted least squares method, which can address the heteroscedascity problem. Model convergence is not guaranteed; whether or not the model converges depends on the data and the choice

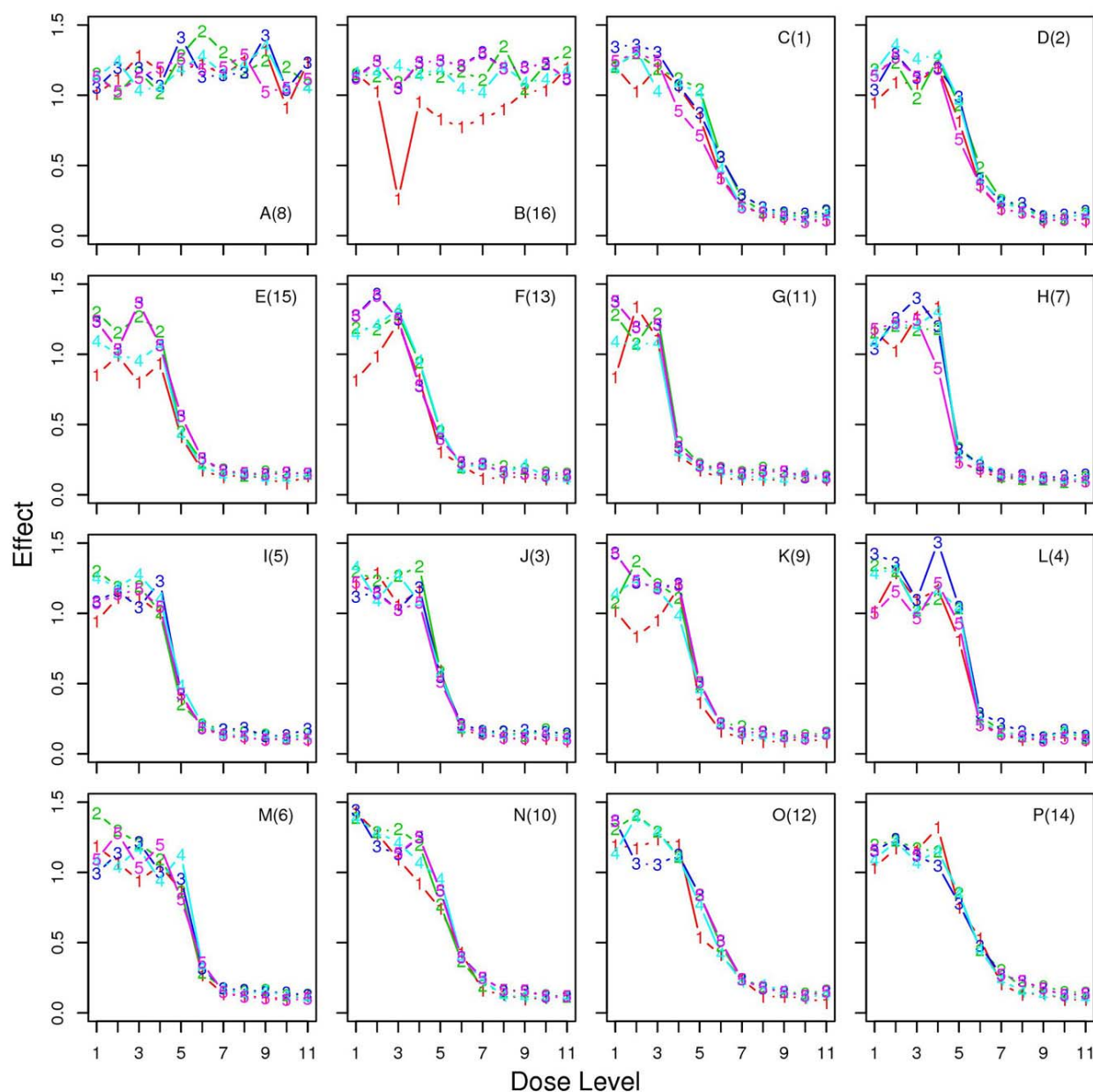


Figure 4. Distribution of the effect versus dose level for curves A through P for the experiment in a low folic acid medium.

of the initial values. We find that PROC NLIN in SAS provides a more comprehensive and robust environment for estimating parameters with nonlinear regression compared to the `nls()` function in S-PLUS/R. It can be useful to apply SAS first to estimate the parameters and then feed the results into S-PLUS/R for further data analysis and production of graphics. Unlike fitting the linearly-transformed median effect model via linear regression, for which a solution can always be found, fitting the E_{max} model via nonlinear regression may result in nonconvergence of the model in some cases. This nonconvergence may indicate aberrant conditions in the data such that the data do not provide adequate information

for model fitting. We had convergence problems with the curves G, H, and K in the low FA experiment. In these cases, there were insufficient data in the middle of the effect range; hence, the parameters could not be estimated reliably. We had to fix the m parameter before we could estimate the other two parameters. From the dose-response curves, we found that TMQ was more potent than AG, and that the drug combination was more potent in the low FA medium than in the high FA medium.

Upon construction of the marginal and combination dose-response curves, we applied the Loewe additivity model to compute the interaction index. We note

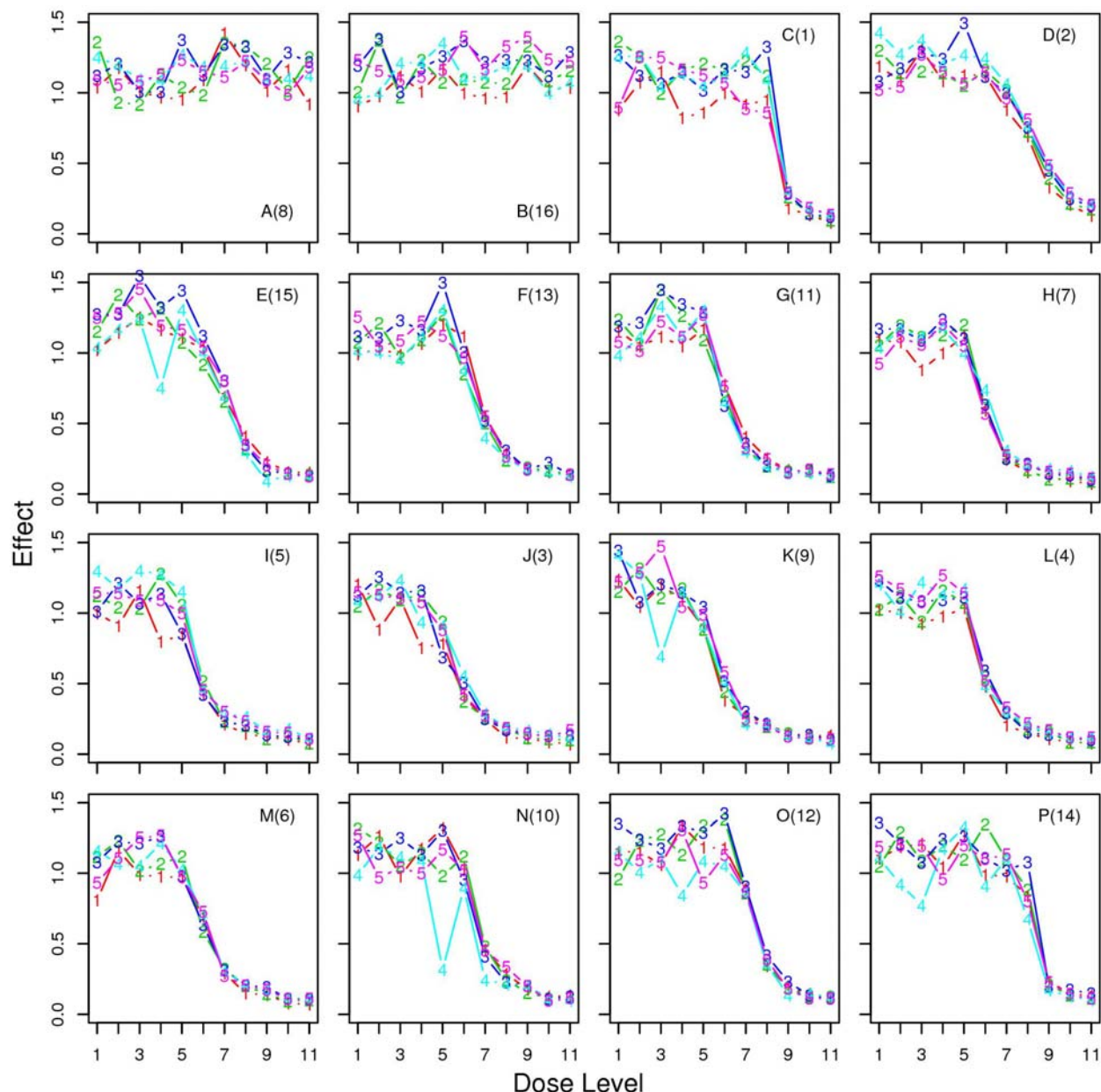


Figure 5. Distribution of the effect versus dose level for curves A through P for the experiment in a high folic acid medium.

that a definition of drug interaction such as the interaction index is model dependent. Additionally, no matter which model is used, based on the definition of the interaction index (7,8), the dose levels used in calculating the interaction index must be translated back to the original units of dose measurement. Under the given model, we found that the drug interaction between TMQ and AG was largely synergistic. Synergy was more clear and evident in the high FA experiment than in the low FA experiment. In addition, synergy was more likely to be observed when a small dose of the more potent drug (TMQ) was added to a large dose of the less potent drug (AG). When a large amount of a

more potent drug is present, adding the less potent drug does not show synergy because the effect is already largely achieved by the more potent drug. In addition, the interval estimation showed that the 95% confidence intervals were wider at the two extremes of the effect, which were closer to 1 or to $1-E_{max}$. This result is consistent with that of many regression settings in which estimation achieves higher precision in the center of the data distribution but lower precision at the extremes.

We have provided a simple, yet useful approach for analyzing drug interaction for combination studies. The interaction index for each fixed dose ratio

Emax model and interaction index for drug interaction

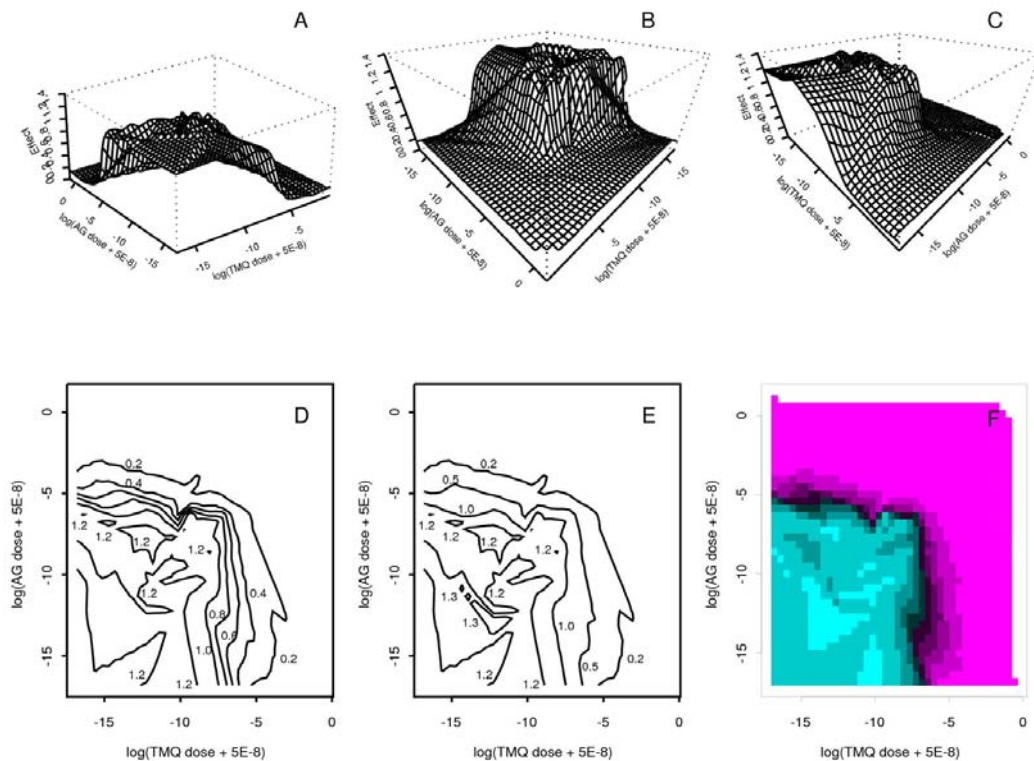


Figure 6. Perspective plots (A, B, C), contour plots (D, E), and image plot (F) for the effect versus logarithm transformed doses of trimetrexate and AG2034 for the experiment in a low folic acid medium.

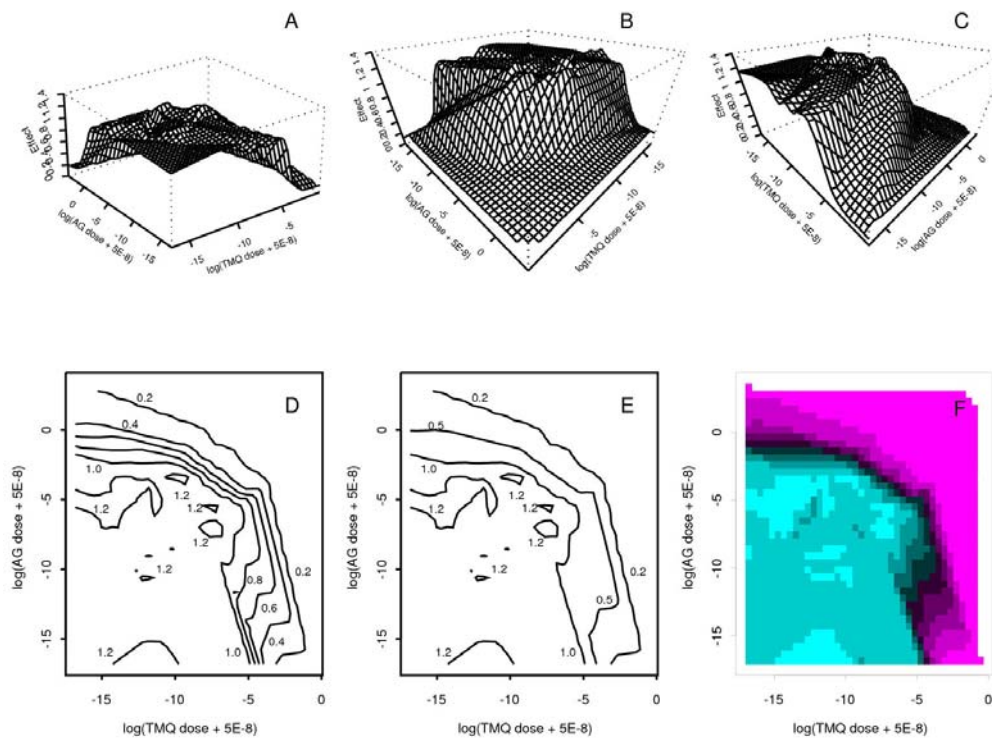


Figure 7. Perspective plots (A, B, C), contour plots (D, E), and image plot (F) for the effect versus logarithmically transformed doses of trimetrexate and AG2034 for the experiment in a high folic acid medium.

Emax model and interaction index for drug interaction

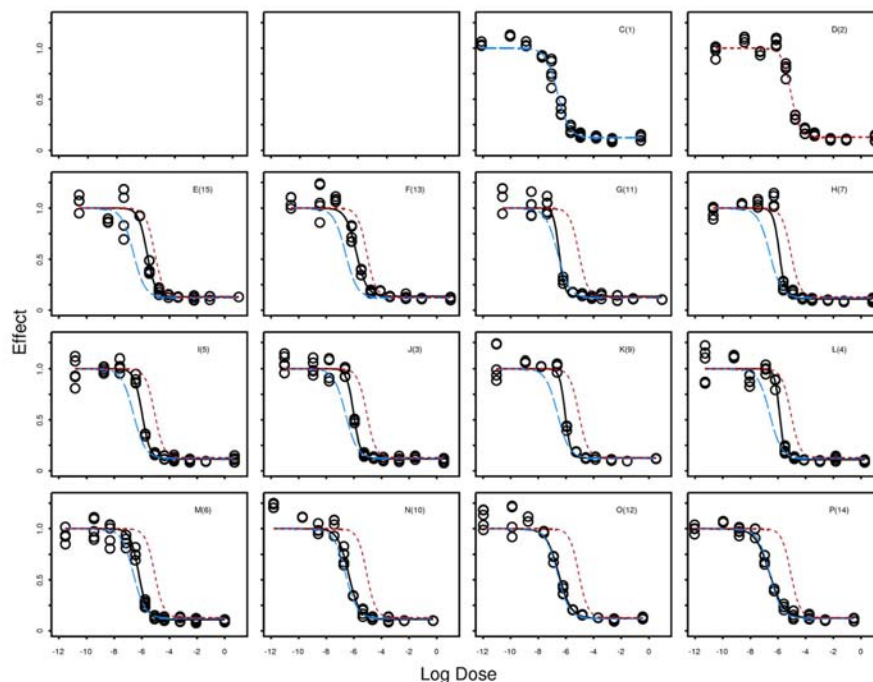


Figure 8. Effect versus logarithmically transformed dose plot for the combination study of trimetrexate and AG2034 in a low folic acid medium. Raw data are shown in open circles. Blue dashed line and red dotted line indicate the fitted marginal dose-response curves for trimetrexate and AG2034, respectively. Black solid line indicates the fitted dose-response curve for the combination of trimetrexate and AG2034.

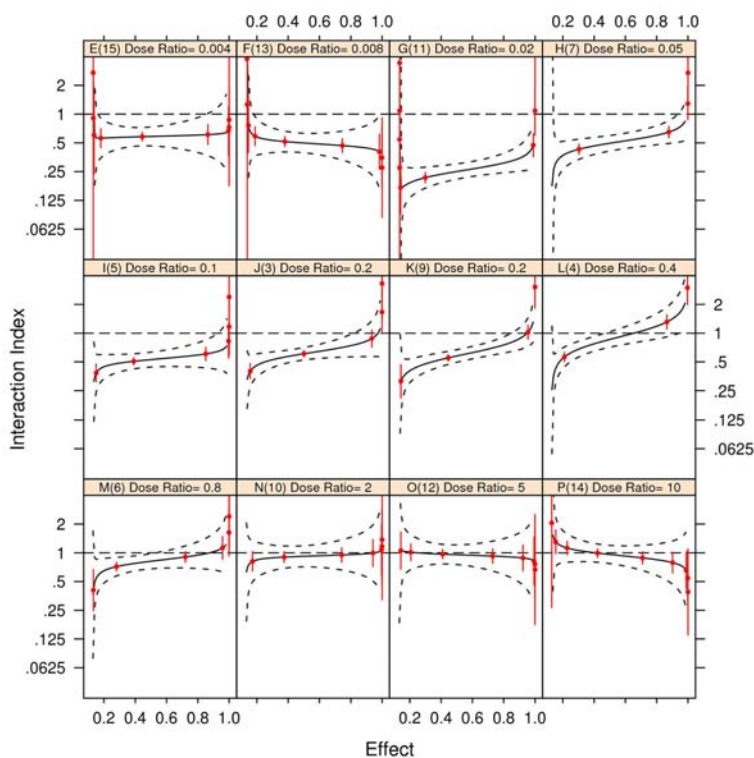


Figure 9. Trellis plot of the estimated interaction index (solid line) and its point-wise 95% confidence interval (red solid lines) and the 95% simultaneous confidence band (dashed lines) for the low folic acid experiment. Estimates at the design points where experiments were conducted are in red. The interaction index is plotted on the logarithmically transformed scale but labeled on the original scale.

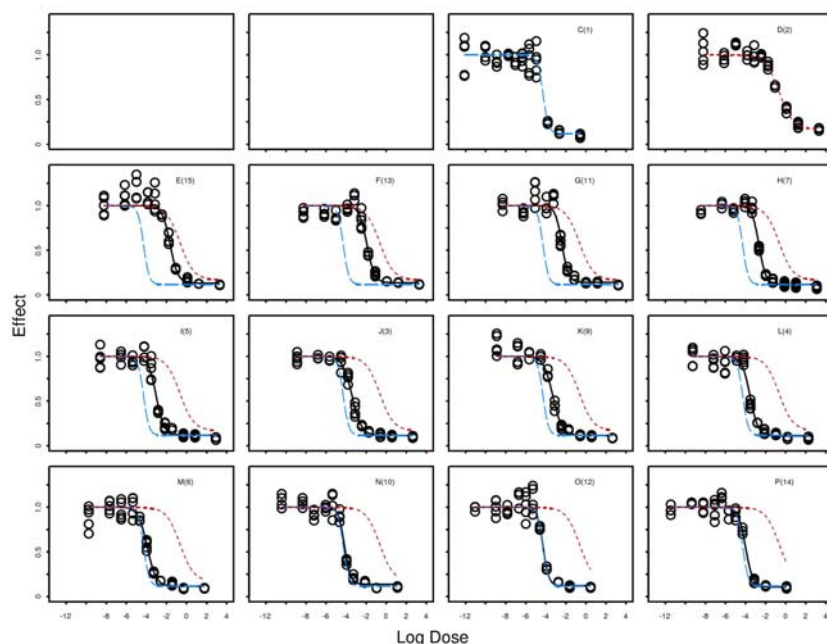


Figure 10. Effect versus logarithmically transformed dose plot for the combination study of trimetrexate and AG2034 in a high folic acid medium. Raw data are shown in open circles. Blue dashed line and red dotted line indicate the fitted marginal dose-response curves for trimetrexate and AG2034, respectively. Black solid line indicates the fitted dose-response curve for the combination of trimetrexate and AG2034.

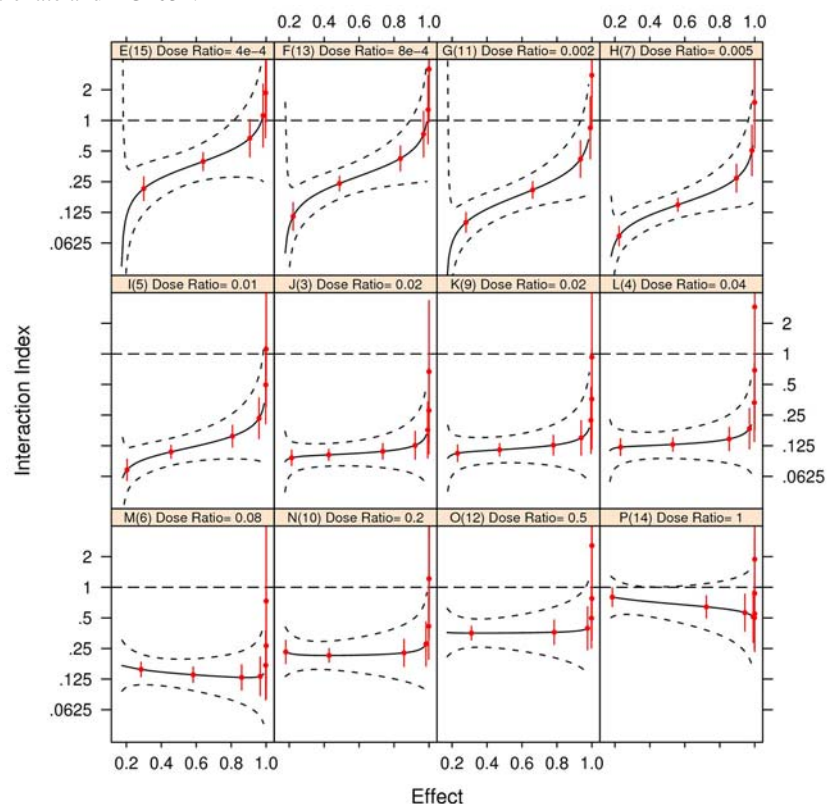


Figure 11. Trellis plot of the estimated interaction index (solid line) and its point-wise 95% confidence interval (red solid lines) and the 95% simultaneous confidence band (dashed lines) for the high folic acid experiment. Estimates at the design points where experiments were conducted are in red. The interaction index is plotted on the logarithmically transformed scale but labeled on the original scale.

is computed and then displayed together using a trellis plot. This method works well for the ray design. Other methods have been proposed to model the entire response surface using the parametric approach (27) or the semiparametric approach (28). The results from applying the semiparametric model are reported in a companion article (29).

9. ACKNOWLEDGEMENT

We thank Dr. William R. Greco at the Roswell Park Cancer Institute for organizing this project comparing rival modern approaches to analyzing combination studies, for supplying the data sets, and for the invitation to present this manuscript. The authors also thank Lee Ann Chastain for her editorial assistance. This work is supported in part by grants W81XWH-05-2-0027 and W81XWH-07-1-0306 from the Department of Defense, and grant CA16672 from the National Cancer Institute. J. Jack Lee's research was supported in part by the John G. & Marie Stella Kenedy Foundation Chair in Cancer Research.

10. REFERENCES

1. R. Chris Rathbun, Staci M. Lockhart, Johnny R. Stephens: Current HIV treatment guidelines - an overview. *Curr Pharm Des* 12, 1045-1063 (2006)
2. Julian R. Molina, Alex A. Adjei, James R. Jett: Advances in chemotherapy of non-small cell lung cancer. *Chest* 130, 1211-1219 (2006)
3. Myaing M. Nyunt, Christopher V. Plowe: Pharmacologic advances in the global control and treatment of malaria: combination therapy and resistance. *Clin Pharmacol and Ther* 82, 601-605 (2007)
4. Kamalesh K. Sankhala, Kyriakos P. Papadopoulos: Future options for imatinib mesilate-resistant tumors. *Expert Opinion Investig Drugs* 16, 1549-1560 (2007)
5. Roberto Bianco, Vincenzo Damiano, Teresa Gelardi, Gennaro Daniele, Fortunato Ciardiello, Giampaolo Tortora: Rational combination of targeted therapies as a strategy to overcome the mechanisms of resistance to inhibitors of EGFR signaling. *Curr Pharm Des* 13, 3358-3367 (2007)
6. Sith Sathornsumetee, David A. Reardon, Annick Desjardins, Jennifer A. Quinn, James J. Vredenburgh, Jeremy N. Rich: Molecularly targeted therapy for malignant glioma. *Cancer* 110, 12-24 (2007)
7. Morris C. Berenbaum: What is synergy? *Pharmacol Rev* 41, 93-141 (1989)
8. William R. Greco, Gregory Bravo, John C. Parsons: The search of synergy: A critical review from a response surface perspective. *Pharmacol Rev* 47(2), 331-385 (1995)
9. Jurgen Suhnel: Parallel dose-response curves in combination experiments. *Bull Math Biol* 60, 197-213 (1998)
10. Ting-Chao Chou: Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol Rev* 58, 621-681 (2006)
11. Ronald J. Tallarida: An overview of drug combination analysis with isobolograms. *J Pharmacol Exp Ther* 319, 1-7 (2006)
12. Ting-Chao Chou, Paul Talalay: Quantitative analysis of dose effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv Enzyme Reg* 22, 27-55 (1984)
13. Stella G. Machado, Grant A. Robinson: A direct, general approach based on isobolograms for assessing the joint action of drugs in pre-clinical experiments. *Stat Med* 13, 2289-2309 (1994)
14. Chris Gennings: On testing for drug/chemical interactions: definitions and inference. *J Biopharm Stat* 10, 457-467 (2000)
15. Kathryn S. Dawson, Walter H. Carter, Jr., Chris Gennings: A statistical test for detecting and characterizing departures from additivity in drug/chemical combinations. *J Agric Biol Environ Stat* 5, 342-359 (2000)
16. J. Jack Lee, Maiying Kong, Gregory D. Ayers, Reuben Lotan: Interaction index and different methods for determining drug interaction in combination therapy. *J Biopharm Stats* 17, 461-480 (2007)
17. J. Jack Lee, Maiying Kong: A confidence interval for interaction index for assessing multiple drug interaction. *Stat Biopharm Res* 1, 4-17, (2009)
18. Helene M. Faessel, Harry K. Slocum, Robert C. Jackson, Theodore J. Boritzki, Youcef M. Rustum, M. G. Nair, and William R. Greco: Super in vitro synergy between inhibitors of dihydrofolate reductase and inhibitors of other folate-requiring enzymes: The critical role of polyglutamylolation. *Cancer Res* 58, 3036-3050 (1998)
19. Naitee Ting: Dose finding in drug development. Springer, NY, 127-145 (2006)
20. John M. Chambers, Trevor J. Hastie: Statistical models in S. Chapman & Hall/CRC Press, Boca Raton FL, 450 – 452 (1992)
21. Sylvie Huet, Anne Bouvier, Marie-Anne Poursat, Emmanuel Jolivet: Statistical tools for nonlinear regression: a practical guide with S-PLUS and R examples, 2nd ed. Springer, NY (2003)
22. SAS/STAT 9.1 User's guide. SAS Institute, Cary, NC (2007)

23. Peter J. Bickel, Kjell A. Doksum: Mathematical statistics: basic ideas and selected topics, Vol I, 2nd ed. *Prentice Hall*, New Jersey pp. 306-314 (2000)
24. Christopher Cox, Guangpeng Ma: Asymptotic confidence bands for generalized nonlinear regression models. *Biometrics* 51, 142-150 (1995)
25. J. Jack Lee, Z. Nora Tu: A versatile one-dimensional distribution plot: the BLiP plot. *Amer Statistician* 51(4), 353-358 (1997)
26. Michael J. Crawley: The R book. *Wiley*, NY (2007)
27. Maiying Kong, J. Jack Lee: A general response surface model with varying relative potency for assessing drug interactions. *Biometrics* 62(4), 986-995 (2006)
28. Maiying Kong, J. Jack Lee: A semiparametric response surface model for assessing drug interactions. *Biometrics* 64, 396-405 (2008)
29. Maiying Kong, J. Jack Lee: Applying E_{max} model and bivariate thin plate splines to assess drug interactions. *Front Biosci* In Press

Abbreviations: AG: AG2034, an inhibitor of the enzyme glycylamide ribonucleotide formyltransferase; ED₅₀: dose required to produce 50% of the maximum effect; E_{max}: maximum effect attributed to the drug; FA: folic acid; II: interaction index; TMQ: Trimetrexate, a lipophilic inhibitor of the enzyme dihydrofolate reductase;

Key Words: additivity, antagonism, confidence interval estimation, E_{max} model, Loewe additivity model, nonlinear regression, synergy, trellis plot

Send correspondence to: J. Jack Lee, Department of Biostatistics, Unit 1411, The University of Texas M. D. Anderson Cancer Center, P. O. Box 301402, Houston, Texas 77230-1402, Tel: 713-794-4158, Fax: 713-563-4242, E-mail: jjlee@mdanderson.org

<http://www.bioscience.org/current/vol2E.htm>

Nrf2 and Keap1 Abnormalities in Non-Small Cell Lung Carcinoma and Association with Clinicopathologic Features

Luisa M. Solis¹, Carmen Behrens², Wenli Dong³, Milind Suraokar², Natalie C. Ozburn², Cesar A. Moran¹, Alejandro H. Corvalan^{1,2}, Shyam Biswal⁸, Stephen G. Swisher⁴, B. Nebiyu Bekele³, John D. Minna^{5,6,7}, David J. Stewart² and Ignacio I. Wistuba^{1,2}

Authors' Affiliations: Departments of ¹Pathology, ²Thoracic/Head and Neck Medical Oncology, ³Biostatistics, and ⁴Thoracic Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas; ⁵Hamon Center for Therapeutic Oncology Research and Departments of ⁶Internal Medicine and ⁷Pharmacology, The University of Texas Southwestern Medical Center, Dallas, Texas; and ⁸Department of Environmental Health Sciences, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland.

Grant information: This study was supported in part by grants from the Department of Defense (W81XWH-07-1-0306 to J.D.S., J.D.M., and I.I.W.), the Specialized Program of Research Excellence in Lung Cancer Grant (P50CA70907 to J.D.M. and I.I.W.; P50CA058184 to S.B.), and the National Cancer Institute (Cancer Center Support Grant CA-16672).

Running title: Keap1/Nrf2 in NSCLC

Abbreviations: Nrf2, nuclear factor erythroid-2 related factor 2; Keap1, Kelch-like ECH-associated protein 1; NSCLC, non-small cell lung carcinoma.

Key words: Nrf2, Keap1, NSCLC

Requests for reprints: Ignacio I. Wistuba, Departments of Pathology and Thoracic/Head and Neck Medical Oncology, Unit 85, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030; phone: 713-563-9184; fax: 713-792-0309; e-mail: iiwistuba@mdanderson.org.

Translational Relevance

Nuclear factor erythroid-2-related factor 2 (Nrf2) is a transcription factor associated with chemotherapy resistance and tumor growth, which is repressed by Kelch-like ECH-associated protein 1 (Keap1). We tested the hypothesis that the abnormal expression of these two proteins correlated with non-small cell lung cancer (NSCLC) patients' outcome and response to adjuvant chemotherapy. We demonstrated that increased Nrf2 expression and decreased Keap1 expression are common abnormalities in NSCLC and are associated with clinical outcome. In our study, abnormal expression of Nrf2 and Keap1 proteins was more common than that of the corresponding gene mutations, suggesting that other mechanisms are involved in the activation of *NFE2L2* and inactivation of *KEAP1*. Nrf2 expression may play a role in response to adjuvant platinum-based chemotherapy in patients with squamous cell carcinoma. Identifying patients with abnormal Nrf2 expression may be important for selection for chemotherapy in NSCLC.

ABSTRACT

Purpose. To understand the role of Nrf2 and Keap1 in NSCLC, we studied their expression in a large series of tumors with annotated clinicopathologic data, including response to platinum-based adjuvant chemotherapy.

Experimental design. We determined the immunohistochemical expression of nuclear Nrf2 and cytoplasmic Keap1 in 304 NSCLCs and its association with patients' clinicopathologic characteristics, and in 89 tumors from patients who received neoadjuvant (n=26) or adjuvant platinum-based chemotherapy (n=63). We evaluated *NFE2L2* and *KEAP1* mutations in 31 tumor specimens.

Results. We detected nuclear Nrf2 expression in 26% of NSCLCs; it was significantly more common in squamous cell carcinomas (38%) than in adenocarcinomas (18%; $P<0.0001$). Low or absent Keap1 expression was detected in 56% of NSCLCs; it was significantly more common in adenocarcinomas (62%) than in squamous cell carcinomas (46%; $P=0.0057$). In NSCLC, mutations of *NFE2L2* and *KEAP1* were very uncommon (2 of 29 and 1 of 31 cases, respectively). In multivariate analysis, Nrf2 expression was associated with worse overall survival ($P=0.0139$; HR=1.75) in NSCLC patients, and low or absent Keap1 expression was associated with worse overall survival ($P=0.0181$; HR=2.09) in squamous cell carcinoma. In univariate analysis, nuclear Nrf2 expression was associated with worse recurrence-free survival in squamous cell carcinoma patients who received adjuvant treatment ($P=0.0410$; HR=3.37).

Conclusions. Increased expression of Nrf2 and decreased expression of Keap1 are common abnormalities in NSCLC and are associated with a poor outcome. Nuclear expression of Nrf2 in malignant lung cancer cells may play a role in resistance to platinum-based treatment in squamous cell carcinoma.

INTRODUCTION

Lung cancer is the most common cause of cancer-related death in the world (1). Non-small cell lung carcinoma (NSCLC), adenocarcinoma, and squamous cell carcinoma are the most common histologic types (85%) (2). Despite intensive research, the prognosis of lung cancer patients remains poor, with a 15% 5-year overall survival (OS) rate (1). For patients with early-stage disease, surgery is the standard treatment (2). Adjuvant chemotherapy has been found to be beneficial for some patients with stage II-IIIa NSCLC (3), whereas for patients with stage IIIB and IV disease, chemotherapy is the standard front-line treatment (3, 4). A combination of drugs that includes a platinum agent is the most common regimen administered for NSCLC (3, 4). However, most tumors either fail to respond due to intrinsic resistance or else develop drug acquired resistance after an initial response to therapy (3).

Nuclear factor erythroid-2-related factor 2 (Nrf2) is a transcription factor that has been suggested to be associated with cancer development and progression, including in NSCLC (5-8). Nrf2 enables the adaptation of normal cells to oxidants and electrophiles generated by harmful exogenous agents and to reactive oxygen species and their secondary metabolites (9). Under homeostatic conditions, Nrf2 is principally repressed by Kelch-like ECH-associated protein 1 (Keap1), which functions as an intracellular redox sensor, targeting Nrf2 for proteosomal degradation. Under oxidant or xenobiotic stress, Keap1 releases Nrf2, which translocates to the nucleus and activates antioxidant response elements and xenobiotics element genes (including NAD(P)H dehydrogenase quinone 1, NQO1), resulting in the protein expression of growth factors and receptors, drug efflux pumps, drug-metabolizing enzymes, heat shock proteins, and various transcription factors (5, 9, 10).

One of the mechanisms involved in NSCLC is the nuclear translocation of Nrf2 due to

loss of Keap1 expression by biallelic inactivation of the gene by mutation and loss of heterozygosity, or promoter methylation (5, 8, 11-13). An alternative mechanism of Nrf2 activation is mutation of the gene *NFE2L2*, which affects the region of exon 2 that codes the Keap1-binding site of Nrf2 (12); these mutations have been detected in 8-11% of NSCLCs, mainly squamous cell carcinoma tumors (12, 14). It has been suggested that abnormalities of the Nrf2/Keap1 pathway that lead to nuclear Nrf2 expression in tumors are an important mechanism to induce platinum-based chemotherapy resistance by promoting tumor cell survival and increasing proliferation (5-8). Recently, it was shown that the inhibition of Nrf2 expression using siRNA augmented carboplatin-induced tumor growth inhibition in an NSCLC xenograft mouse model (8).

To date, no comprehensive analysis has been performed of Nrf2 and Keap1 expression and associated genetic abnormalities in NSCLC, and no studies have determined the relationship between Nrf2 expression and clinical outcome after treatment with platinum-based adjuvant chemotherapy. Therefore, in this retrospective study, we characterized the expression of these two proteins in a large series of NSCLC tissue specimens with annotated clinicopathologic characteristics, including outcome, determined the frequency of exon 2 *NFE2L2* and exon 2-5 *KEAP1* mutations and evaluated the relationship between nuclear Nrf2 expression and outcome in patients treated with platinum-based adjuvant chemotherapy. Because in NSCLC the presence of *EGFR* and *KRAS* mutation has been associated to tumor's response to chemotherapy (15), we also investigated in adenocarcinoma tumors the association between both genes mutation status and Nrf2 and Keap1 expressions.

MATERIALS AND METHODS

Nrf2 and Keap1 Western blot analysis in cell lines. The human NSCLC cell lines A549 and H460 (with known downregulation of Keap1 protein) (5), H1993, and an SV40-transformed human bronchial epithelial cell line, BEAS2B, were evaluated for nuclear and cytoplasmic expression of Nrf2 and Keap1 proteins by Western blot analysis. Nuclear and cytoplasmic protein extracts of these cell lines were obtained using NE-PER nuclear extraction reagents (Pierce, Rockford, IL). The lung cancer cell lines were provided by Dr. John Minna's laboratory (Dallas, TX) and they were authenticated by testing them using the PowerPlex^R 2.1 system (Promega, Madison, WI). Protein concentrations were estimated using the Bradford assay (Bio-Rad, Hercules, CA). For Western blot analysis, 25 µg of cell line protein from nuclear and cytoplasmic extracts were loaded in each lane, run on a NuPAGE 4%-12% Bis-Tris gel (Invitrogen, Carlsbad, CA), and transferred onto a nitrocellulose membrane. After being blocked with 5% non-fat milk, the blots were exposed to rabbit primary antibody against Nrf2 (dilution 1:500, clone H300; Santa Cruz Biotechnologies, Santa Cruz, CA) and Keap1 (dilution 1:600; Proteintech, Chicago, IL), followed by anti-rabbit secondary antibody. The signals were detected using SuperSignal West Pico chemiluminescent substrate (Pierce, Rockford, IL). β-actin and poly(ADP-ribose) polymerase (dilution 1:100, Cell Signaling Technologies, Danver, MA) were used as the controls. The Western blot analysis was performed in triplicate.

Case selection for immunohistochemical analysis. To determine the expression of Nrf2 and Keap1 in primary NSCLCs, we selected archived, formalin-fixed, paraffin-embedded (FFPE) tumor tissue samples from surgically resected lung cancer specimens from the Lung Cancer specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, Texas). This study was approved by the M. D. Anderson

Cancer Center institutional review board. Tumor tissues were histologically analyzed and classified using the 2004 World Health Organization classification system (16). These samples were used to evaluate the immunohistochemical expression of Nrf2 and Keap1 in both a tissue microarray (TMA) and whole tissue sections.

For the TMA, we used 304 tumor tissue samples collected between 1997 and 2003, including 190 adenocarcinomas and 114 squamous cell carcinomas. The cases were selected based on the availability of FFPE tissue blocks with enough tumor tissue for TMA construction. These samples were placed in a TMA, using three 1-mm-diameter cores that included tissue from the center, intermediate, and peripheral areas of the tumor, as previously described (17). Detailed clinicopathologic information, including demographics, performance status (based on Eastern Cooperative Oncology Group, ECOG, scale), smoking history (never, former, or current), pathologic TNM stage (I-IV) (Table 1), recurrence-free survival (RFS) and overall survival (OS) duration, were available for most cases. To determine the heterogeneity of nuclear Nrf2 expression in NSCLC tissues, we evaluated whole tumor tissue sections from 36 tumors, including 18 adenocarcinomas and 18 squamous cell carcinomas; 19 of these cases expressed nuclear Nrf2. Thirty of these cases were also examined for the expression of NQO1. The whole tumor histology sections consisted of 1 to 2 cm diameter tumor specimens with adjacent normal lung tissue.

To determine the expression of Nrf2 and Keap1 proteins in NSCLC after chemotherapy, we evaluated 26 tumor tissues from patients who had undergone neoadjuvant platinum-based chemotherapy. The chemotherapy regimens included carboplatin with paclitaxel (n=21) or cisplatin with etoposide (n=4) or docetaxel (n=1). We also determined the association between Nrf2 and Keap1 expression and histologic parameters associated with chemotherapy effects in

tumor tissues, including the percentages of tumor necrosis, fibrosis, and viable malignant cells.

To determine the relationship between nuclear Nrf2 expression and outcome after adjuvant chemotherapy, we selected 122 NSCLC tumors, 63 from patients who had undergone adjuvant platinum-based chemotherapy, and a similar number of patients who had not undergone any adjuvant therapy (n=59) (Supplementary Table 1). The chemotherapy regimens included carboplatin with docetaxel (n=9), gemcitabine (n=9), paclitaxel (n=32), or cisplatin (n=1), either alone or with pemetrexed (n=4), docetaxel (n=7), etoposide (n=1).

Immunohistochemical analysis of cell lines and tissue specimens. An immunohistochemical analysis was performed using commercially available antibodies against Nrf2 (dilution 1:200, clone H300; Santa Cruz Biotechnologies), Keap1 (dilution 1:25; Proteintech), and NQO1 (dilution 1:1000, clone A180; Novus Biological, Littleton, CO). Immunohistochemical staining was performed using an automated stainer (Dako, Inc., Carpinteria, CA) with 5- μ M-thick sections from FFPE tissues. Tissue sections were deparaffinized and hydrated, and antigen retrieval was performed in pH 6.0 citrate buffer in a decloaking chamber (121°C x 30 seconds, 90°C x 10 seconds) and washed on Tris buffer. Peroxide blocking was performed at ambient temperature for 30 minutes with 3% H₂O₂ in distilled water for Nrf2 and methanol for Keap1 and NQO1. The slides were incubated with primary antibody at ambient temperature and washed with Tris buffer, followed by incubation with biotin-labeled secondary antibody for 15 minutes and streptavidin peroxidase for 15 minutes (LSAB system, Dako) for Nrf2 and Envision Dual-Link system-horseradish peroxidase (Dako) for 30 minutes for Keap1 and NQO1. Staining was developed with 0.5% 3,3'-diaminobenzidine, freshly prepared with imidazole-HCl buffer, pH 7.5, containing hydrogen peroxide and an antimicrobial agent (Dako) for 5 minutes and then

counterstained with hematoxylin, dehydrated, and mounted.

To determine the association between Nrf2 and Keap1 expression on Western blot and immunohistochemical analyses, we prepared FFPE cell pellets from 4 cells lines (A549, H460, H1993, and BEAS2B). For the immunohistochemical analysis, the pellets were used as positive controls. As negative controls, we used positive control sections, replacing the primary antibody with universal negative control anti-rabbit (Dako).

Immunohistochemical expression was quantified jointly by 2 pathologists (L.S. and I.W.). Nuclear Nrf2, cytoplasmic Keap1 and cytoplasmic NQO1 expressions were quantified using a 4-value intensity score (0, 1+, 2+, or 3+) and the percentage (0%-100%) of the extent of reactivity. An immunohistochemical expression score was obtained by multiplying the intensity and reactivity extension values (range, 0-300), and these expression scores were used to determine expression levels. Positive nuclear Nrf2 expression was defined as a score >0. Low or absent cytoplasmic Keap1 expression was defined as a score <150, which represents the mean of expression for all NSCLC TMA cases. High cytoplasmic NQO1 expression was defined as a score >130, which represents the median expression for NSCLCs evaluated using whole tissue sections.

***EGFR* and *KRAS* mutation analyses.** Exons 18–21 of *EGFR* and exon 1 of *KRAS* were polymerase chain reaction (PCR)-amplified using intron-based primers, as described previously (18, 19), and DNA specimens were extracted from microdissected FFPE tissue. All PCR products were directly sequenced using the PRISM dye terminator cycle-sequencing method (Applied Biosystems, Foster City, CA). All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and were sequenced in both

directions, as reported previously (18, 19).

***NFE2L2* and *KEAP1* mutation analysis in tumor specimens.** To determine the mutation status of *NFE2L2* (exon 2) and *KEAP1* (exons 2 to 5) genes, we selected 31 NSCLC tumors from the TMA set for which DNA extracted from fresh tumor tissue was available. The cases included 20 tumors (9 adenocarcinomas and 11 squamous cell carcinomas) with nuclear Nrf2 expression on immunohistochemical analysis and 11 (8 adenocarcinomas and 3 squamous cell carcinomas) without nuclear Nrf2 expression. The mutation analysis was performed using direct sequencing after PCR amplification of *NFE2L2* and *KEAP1* genes. For *NFE2L2*, intron-based PCR primers (forward 5'-CCACCATCAACAGTGGCATA-3'; reverse 5'-AGGCAAAGCTGGAAGTCAAA-3') for exon 2 were designed using Primer3 software (<http://frodo.wi.mit.edu/>) and synthesized by Sigma-Aldrich (St. Louis, MO). The PCR cycling conditions were 94°C (15 minutes) for 1 cycle; 94°C (30 seconds), 58°C (45 seconds), and 72°C (1 minute) for 45 cycles; and a final extension of 72°C (5 minutes). For the *KEAP1* gene, we analyzed exons 2-5 using intron-based PCR primer sequences, as previously described (5). All PCR products were directly sequenced using the Applied Biosystems PRISM Aldrich (St. Louis, MO) dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications and sequenced in both directions.

Statistical analysis. The clinicopathologic data were summarized using descriptive statistics and frequency tabulations. Wilcoxon rank-sum and Kruskal-Wallis tests were used to compare biomarker expression among different prognostic factor levels. We determined the association between 5-year RFS and OS rates and Nrf2 and Keap1 expression in NSCLC patients with stage

I or II disease who had not undergone adjuvant or neoadjuvant chemotherapy and in patients with stage I-IIIB disease who had and had not undergone platinum-based adjuvant therapy. RFS was defined as the time from surgery to recurrence or the end of the study, and OS was defined as the time from surgery to death or the end of the study. The cut-off for nuclear Nrf2 expression was a score >0 , which was defined as positive staining; and the cut-off for cytoplasmic Keap1 was 150, which represents the mean score. Survival curves were estimated using the Kaplan-Meier method. Univariate and multivariate Cox proportional hazards models were used to assess the effects of covariates on RFS and OS rates. Two-sided *P* values <0.05 were considered statistically significant. All analyses were conducted using SAS (v 9.1, Cary, NC) and S-plus (v 8.0, Seattle, WA) software.

RESULTS

Nrf2 and Keap1 protein expression in cell lines by Western blot analysis and validation of immunohistochemical results. We found on Western blot analysis that Nrf2 protein levels were higher in the nucleus of NSCLC cell lines A549 and H460 than in the cytoplasm (Supplementary Fig. 1). The BEAS2B bronchial epithelial cell line had a significantly lower Nrf2 expression level, both in the nucleus and cytoplasm, than did the malignant cell lines. These findings are consistent with previously reported data (5). In contrast, BEAS2B cells had significantly higher Keap1 expression levels in the nucleus and cytoplasm than did the NSCLC cell lines evaluated, and in all cells, Keap1 was expressed mostly in the cytoplasm (Supplementary Fig. 1). The cell line H1993 had a similar Nrf2 and Keap1 expression pattern to that of BEAS2B on Western blot analysis, which is consistent with H1993 cell line being heterozygous for *KEAP1* mutation (5). We evaluated the immunohistochemical expression of Nrf2 and Keap1 in FFPE cell pellets and

found similar expression patterns to those found on Western blot analysis (Supplementary Fig. 1). In FFPE tumor specimens, in malignant tumor cells, we found both nuclear and cytoplasmic expression of Nrf2 and exclusively cytoplasmic expression of Keap1 (Fig. 1). For the study of the expression of these markers in NSCLC TMAs and whole section tissue specimens, we focused on nuclear Nrf2 expression because this is the subcellular location where is considered to be biologically active (5), and on cytoplasmic Keap1 expression because it was the only expression detected in malignant tumor cells in the FFPE tissue specimens.

Nrf2 and Keap1 immunohistochemical expression in NSCLC TMA and association with clinicopathologic and genetic features. We determined nuclear Nrf2 and cytoplasmic Keap1 protein expression in 304 tumors in TMAs using levels and scores of expression. Positive nuclear Nrf2 expression (score >0) was detected in 26% of NSCLCs, and the frequency of positive cases was significantly ($P < 0.001$) higher in squamous cell carcinomas (38%) than in adenocarcinomas (18%) (Table 2). In most positive tumors (49/77, 64%), the Nrf2 nuclear expression was mild, and in the remaining cases was moderate to strong. On the other hand, negative or low levels of cytoplasmic Keap1 expression (score <150) were observed in 56% of NSCLCs. The frequency of low or negative Keap1 expression was significantly higher in adenocarcinomas (62%) than in squamous cell carcinomas (46%) ($P = 0.0057$). Overall, the expression of nuclear Nrf2 was statistically associated ($P = 0.0041$; $r = 0.17$) with higher cytoplasmic Keap1 expression in all NSCLCs. However, we identified a subset (39 of 295 [13%]) of tumors with positive Nrf2 and low or absent Keap1 expression, including squamous cell carcinomas (19 of 111 [17%]) and adenocarcinomas (20 of 184 [11%]). Nrf2 and Keap1 expression was not associated with sex, smoking history, or pathologic tumor stage (stages I or II

vs. III or IV; data not shown).

Among adenocarcinomas, *EGFR* and *KRAS* mutations were detected in 23 of 172 (13%) and 28 of 171 (16%) cases, respectively. Interestingly, no *EGFR* mutant adenocarcinomas expressed Nrf2, whereas 21% of *EGFR* wild-type tumors expressed nuclear Nrf2; this difference was statistically significant ($P=0.009$). Although *EGFR* mutations were significantly higher in tumors from non-smokers (13/40, 33%) compared with smokers (10/130, 8%), the distribution of smoking status of the 23 patients with *EGFR* mutation and lack of Nrf2 expression was similar: 10 (43%) smokers and 13 non-smokers (57%). There was no association between cytoplasmic Keap1 expression and *EGFR* mutation status. No relationship was found between the expression of these 2 markers and tumors with *KRAS* mutation.

***NFE2L2* and *KEAP1* mutation analysis.** In NSCLC tumor, we studied mutation status of both genes at sites (*NFE2L2*, exon 2; and, *KEAP1*, exons 2 to 5) to be previously reported as mutated in lung cancer (5, 12, 13). To correlate mutation status of both genes with activation of Nrf2, we selected 20 tumors with DNA-extracted from fresh tissue available, and with nuclear Nrf2 immunohistochemical expression; as controls, we used 11 tumors lacking nuclear Nrf2 expression. *NFE2L2* mutation was found in 2 of 29 tumors successfully examined. Both mutations were located in codon 28 (ACA to ATA; substitution of threonine for isoleucine) and codon 79 (CAG to GAG; substitution of glutamic acid for glutamine) of exon 2. These 2 cases corresponded to squamous cell carcinomas with nuclear Nrf2 expression and high cytoplasmic Keap1 expression. *KEAP1* mutation (exons 2 to 5) was detected in only 1 of 31 tumors examined. This was a non-sense mutation (TAC to TAA; substitution of tyrosine for stop codon) in codon 537 (exon 5) in a squamous cell carcinoma with nuclear Nrf2 expression and low

cytoplasmic Keap1 expression.

Nrf2, Keap1 and NQO1 immunohistochemical expression in NSCLC tumors and corresponding normal adjacent tissues using whole tissue section analysis. To determine the heterogeneity of nuclear Nrf2 and cytoplasmic Keap1 expression in tumor tissues, we evaluated the expression of both markers using whole histologic sections obtained from 36 NSCLCs (18 adenocarcinomas and 18 squamous cell carcinomas) included in the TMA, 19 of which were Nrf2 positive and 17 of which were negative. Nuclear Nrf2 expression was highly heterogeneous throughout the positive tumors, with 14 (78%) expressing nuclear Nrf2 in 5%-30% of malignant cells. Of interest, only 1 of the 17 tumors that was Nrf2 negative on the TMA analysis was positive on whole histologic section examination.

To determine the biological effect of Nrf2 expression in NSCLC malignant cells, we studied the correlation of expression of nuclear Nrf2 with the immunohistochemical protein expression of NQO1, a gene transcriptionally regulated by Nrf2 (5, 10), using a subset (n=30) NSCLC with whole tissue sections available. Of interest, 12 of 16 (75%) nuclear Nrf2 positive cases expressed high levels of cytoplasmic NQO1, while only 3 of 12 nuclear Nrf2 negative tumors expressed this protein at high levels. Similarly, the tumors expressing nuclear Nrf2 had a significantly ($P=0.0211$) higher NQO1 expression score (mean 176.3) compared with the Nrf2 negative tumors (mean 92.9).

On the other hand, we found that regardless of the intensity of expression, cytoplasmic Keap1 was homogeneously expressed throughout the tumors. Only 4 of 36 (11%) tumors showed small distinct areas of malignant cells (defined as 10%-30% of the tumor section) lacking cytoplasmic Keap1 expression.

Nrf2 expression was detected in the nucleus of normal bronchial epithelia adjacent tumors in 6 of 127 (5%) chemotherapy-naïve NSCLCs on whole histologic section, including 19 cases with Nrf2-positive tumors. The 6 cases with nuclear Nrf2 expression in normal epithelium did not express the marker in their corresponding tumors. As expected, Keap1 cytoplasmic expression was found in normal bronchial epithelium in all 25 chemotherapy-naïve NSCLCs evaluated. Similar cytoplasmic Keap1 expression scores were detected in the tumors (mean score=126.0, SD=81.8) and corresponding normal bronchial epithelia (mean score=130.0, SD=59.2).

Association between Nrf2 and Keap1 expression and NSCLC patient outcome using TMA specimens. We determined the association between nuclear Nrf2 and cytoplasmic Keap1 expression and RFS and OS rates in patients with stage I and II NSCLC who had not undergone neoadjuvant or adjuvant treatment. In patients with NSCLC (n=235), positive nuclear Nrf2 expression (score >0) was associated with worse 5-year RFS and OS on univariate analysis and a worse OS on multivariate Cox model analysis ($P=0.0139$; hazard ratio [HR]=1.75; 95% confidence interval [95% CI], 1.12-2.73) when adjusted for age at surgery, smoking history, and pathologic stage (Table 3)(Fig. 2A and 2B). No association was found between nuclear Nrf2 expression and outcome by histologic tumor type (Supplementary Fig. 2).

No association was found between Keap1 expression and patients' outcome for all NSCLCs (Supplementary Fig. 3). Then, we examined the effect of Keap1 in the outcome of patients by individual histologic tumor types. We found that negative and low cytoplasmic Keap1 expression (score <150) was associated with worse 5-year RFS and OS in patients with squamous cell carcinomas on univariate analysis and a worse OS on multivariate Cox model

analysis ($P=0.0181$; HR=2.09; 95% CI, 1.13-3.84) when adjusted for age at surgery and pathologic stage (Table 3) (Fig. 2C and 2D). No association was found between Keap1 expression and outcome in patients with adenocarcinoma (Supplementary Fig. 3).

The subset (39 of 295 [13%]) of NSCLCs tumors with both positive Nrf2 and low or absent Keap1 expression was significantly associated with worse OS ($P=0.0111$; HR=1.97; 95% CI, 1.17-3.31) and RFS ($P=0.0325$; HR=1.69; 95% CI, 1.04-2.73) on the univariate and multivariate analysis when adjusted for age at surgery, pathologic stage and smoking history (Table 3) (Fig. 2E and 2F).

Nrf2 and Keap1 expression in neoadjuvant therapy-treated NSCLC. To investigate if neoadjuvant platinum-based chemotherapy leads to increased nuclear Nrf2 expression in NSCLC tumors, we determined the expression of Nrf2 and Keap1 in 26 surgically resected tumors from patients who had undergone platinum-based chemotherapy before surgery. Nuclear Nrf2 expression (score >0) was found in 7 (27%) tumors (Table 2), but negative or low Keap1 expression (score <150) was detected in 18 (69%). The expression of these markers was not associated with the histologic effects of chemotherapy in tumor tissues, including percentages of tumor necrosis, fibrosis, and viable malignant cells (data not shown).

Association between Nrf2 expression and outcome in patients treated with adjuvant chemotherapy. To evaluate the role of Nrf2 in response to platinum-based chemotherapy, we determined nuclear Nrf2 expression in whole histologic sections of 122 surgically resected NSCLC tumors (stages I to IIIB) and the RFS and OS rates. These samples were obtained from 63 patients who had undergone adjuvant chemotherapy (35 adenocarcinomas and 28 squamous

cell carcinomas) and 59 patients who had not undergone adjuvant treatment (27 adenocarcinomas and 32 squamous cell carcinomas). Overall nuclear Nrf2 expression was detected in 47 (39%) of these NSCLCs (Table 2).

Because the significant difference observed in the nuclear Nrf2 expression frequency between squamous cell carcinomas and adenocarcinomas examined in the TMA specimens, we evaluated the predictive effect of Nrf2 expression in adjuvant-treated patients in both tumors histologies separately. In patients with squamous cell carcinoma who had undergone adjuvant treatment, nuclear Nrf2 expression (score >0) was associated with worse RFS on univariate Cox model analysis ($P=0.0410$; HR=3.37; 95% CI, 1.05-10.81) (Supplementary Fig. 4); however, this association was not significant on multivariate analysis ($P = 0.0618$, HR = 3.11, 95% CI = 0.95 – 10.20). In the same tumor type, nuclear Nrf2 showed a trend towards worse OS in patients who received adjuvant chemotherapy; this association was not significant on univariate Cox model analysis ($P=0.0590$; HR=3.52; 95% CI, 0.95-13.04) and multivariate analysis ($P = 0.092$). In patients with squamous cell carcinomas who had not undergone adjuvant treatment, Nrf2 expression was not associated with the RFS or OS rate (Supplementary Fig. 4).

In patients with adenocarcinoma who had not undergone adjuvant therapy, Nrf2 was statistically associated with RFS ($P = 0.0092$) in univariate analysis, although no association was found between Nrf2 expression and the RFS or OS rate on the multivariate analysis. Of interest, in patients with adenocarcinoma the nuclear expression of Nrf2 did not associate with outcome in patients who received adjuvant chemotherapy.

DISCUSSION

In lung cancer, Nrf2 activation in malignant cells has been associated with tumor

progression and chemotherapy resistance (8, 11, 20-22). High levels of nuclear Nrf2 facilitate cancer cell growth and cell survival as a result of the transactivation of cytoprotective genes (8, 20, 21). In NSCLC, the overexpression of nuclear Nrf2 is principally attributable to genetic and epigenetic alterations and the loss of function of its repressor, Keap1 (5, 21, 23). In NSCLC, *KEAP1* mutations at exons 2-6 have been detected in 50% (n=12) of cell lines (5) and 8% (n=65) and 19% (n=54) of tumors (5, 21). Promoter hypermethylation of *KEAP1* was found in all lung cancer cell lines (3 of 3) and tumor tissues (5 of 5) evaluated (11). In addition, mutations of the Nrf2 gene, *NFE2L2*, that affect the region that codes for Keap1 binding sites have been proposed as an alternative mechanism of Nrf2 activation in lung cancer (12), but Nrf2 mutations have only been found in 2% (2 of 85) of lung cancer cell lines and 8-11% of lung cancer tumor specimens, principally in patients with a smoking history and squamous cell carcinoma histologic tumor type (12, 14).. Despite all these recent findings, the characteristics of NSCLC tumors with Nrf2 activation and loss of Keap1 expression and the role of Nrf2 and Keap1 in the response to platinum-based chemotherapy are not well understood.

We performed a comprehensive immunohistochemical analysis of Nrf2 and Keap1 expression in FFPE NSCLC tumors. We used a validated method in which we tested both antibodies in a panel of NSCLC cell lines with known Nrf2 and Keap1 expression (5) by Western blot analysis and immunohistochemical analysis of FFPE cell pellets. The protein expression pattern and subcellular localization of Nrf2 and Keap1 on Western blot analysis in our cell lines was consistent with previously reported data (5), and the expression found on Western blot analysis was consistent with that in cell line pellets.

In our study of TMA specimens, nuclear Nrf2 was expressed in a subset of NSCLCs (26%), more commonly in squamous cell carcinomas (38%) than in adenocarcinomas (18%). To

our knowledge, only one report exists of the immunohistochemical expression of Nrf2 in lung cancer; evaluations of whole histologic sections from stage I NSCLCs revealed that 62% (55 of 89) of tumors expressed nuclear Nrf2 (24). Although we were not able to obtain detailed information from this report (24), this study seems to differ significantly from ours in the immunohistochemical method used to evaluate Nrf2 expression in tissue specimens, including the quantification method. One intriguing observation of our study is that the frequency of nuclear Nrf2 expression was significantly higher in *EGFR* wild-type (21%) adenocarcinomas than in mutant (0%) tumors, and this association was independent of patients smoking history. Although the number of *EGFR* mutant tumors evaluated was relatively small, this finding is of potential interest. To the best of our knowledge, this association has not been previously reported, and it warrants further study. Our observation concurs with the finding that *EGFR* mutations associates with better survival in advanced NSCLC patients treated with chemotherapy (carboplatin and paclitaxel) with and without an EGFR tyrosine kinase inhibitor (erlotonib) (15). We speculate that the lack of nuclear Nrf2 expression in *EGFR* mutant NSCLCs may contribute to the benefit of administering platinum-based chemotherapy.

Nrf2 can be activated by several mechanisms, including mutations of the gene (*NFE2L2*) affecting the Keap1-binding site (12, 14). In our NSCLC samples, exon 2 *NFE2L2* mutations, which code for the Keap1-binding site region of the Nrf2 protein, were infrequently detected, occurring in 2 of 29 (7%) tumors examined. Interestingly, both tumors were squamous cell carcinomas that expressed nuclear Nrf2 and had high Keap1 cytoplasmic expression levels. Others *NFE2L2* mutation analyses in lung cancer has been conducted; they reported that 8-11% of NSCLCs had mutations on exon 2, including both mutations detected in our study, and they associated with tumors' squamous cell carcinoma histology and patients' smoking history (12,

14).

Our analysis of whole histologic NSCLC sections demonstrated that nuclear Nrf2 expression in tumor tissues was heterogeneous and involved a small percentage (5% to 30%) of malignant cells. Despite the heterogeneity of nuclear Nrf2 expression in tumor cells, we found an association between Nrf2 expression in whole histologic sections and corresponding TMA tissue cores in 36 cases examined. The correlation observed between the expressions of nuclear Nrf2 with cytoplasmic NQO1 protein in a subset of our NSCLC, suggest that Nrf2 is biologically active in the nucleus of the malignant cells. *NQO1* gene is transcriptionally regulated by Nrf2, and, as expected, higher levels of NQO1 protein expression was observed in tumors showing nuclear Nrf2 compared with tumor lacking nuclear Nrf2 (5, 10). Nuclear Nrf2 expression has been reported at different frequencies (range, 54% to 92%) in other epithelial tumors, including squamous cell carcinoma of the head and neck (25) and gallbladder carcinoma (13).

Keap1 is the principal cytoplasmic repressor of Nrf2 (23, 26-28). Ours is the first reported study to determine the frequency of low or absent cytoplasmic Keap1 expression in NSCLC and its association with tumors' clinicopathologic characteristics. Low or absent Keap1 expression was common in NSCLC (56%), mainly in adenocarcinomas. However, we identified only 1 *KEAP1* mutation (exons 2-5) in 31 tumors examined, including 20 with nuclear Nrf2 expression, suggesting that *KEAP1* mutation is not the main mechanism of protein loss or reduction. Our findings differ from those of previous publications that reported *KEAP1* mutations in 8% and 19% of two NSCLC cohorts, predominantly adenocarcinomas (26% and 30%) (5, 21). Our findings of a positively significant correlation between the expression of nuclear Nrf2 and cytoplasmic Keap1, and that only 13% of tumors in our study had low or absent Keap1 and nuclear Nrf2 expression, suggest that other mechanisms, not associated with Keap1

inactivation, promote Nrf2 nuclear localization and subsequent activation. There are few alternative mechanism proposed to activate Nrf2, including phosphorylation of Nrf2 protein by a number of protein kinases, including protein kinase C, extra-cellular regulated kinase, Jun N-terminal kinase and phosphatidylinositol kinase (7). Additionally, there is recent evidence that the presence of certain protein motifs determine Nrf2 subcellular localization (29). The Nrf2 protein has the nuclear export signal (NES) motif, which transports the protein from nucleus to cytoplasm; as well as the nuclear localization signal (NLS) motif, which transports the protein from cytoplasm to nucleus. It has been suggested that the net result of these two driving forces is important to regulate Nrf2 subcellular localization independently of its interaction with Keap1 (29).

Nuclear Nrf2 expression in all NSCLC patients, low or absent Keap1 expression in patients with squamous cell carcinoma as well as the subset of NSCLC with both nuclear Nrf2 and low or absent Keap1 were associated with poor outcome. Regardless of the mechanism that leads to nuclear Nrf2 activation in tumor cells, there is evidence that this phenomenon promotes cell survival in malignant cells (8, 12, 13, 21) and may explain the low RFS and OS rates in our NSCLC patients who had undergone surgical resection with curative intent. Our finding of poor survival in patients with low or absent Keap1 expression suggests that inactivation of this putative tumor suppressor gene affects the growth and progression of tumors by mechanisms that are not mediated by Nrf2. One of those unknown mechanisms could involve other Keap1 binding proteins that have antiapoptotic and proliferative functions (30-32), including phosphoglycerate mutase family member 5 (31), the nuclear oncoprotein prothymosin α (30), and fetal Alz-50 reactive clone 1 protein (32). The association between Keap1 and patient outcome has not been previously reported in human epithelial tumors, except in one study of

renal cell carcinoma that showed that Keap1 overexpression was associated with more advanced tumor stage and poor overall survival (33).

It has been suggested that abnormalities in the Nrf2/Keap1 pathway that lead to Nrf2 overexpression in tumors induce platinum-based chemotherapy resistance by promoting tumor cell survival and increasing proliferation (5-8). In NSCLC cell lines, the upregulation of Nrf2's downstream genes plays a critical role in platinum chemoresistance, mainly because of increased transcription of the genes involved in drug efflux systems and of antioxidant proteins, including glutathione, thioredoxins and NQO1 (20, 24, 34). Furthermore, studies in *in vitro* and *in vivo* NSCLC models have shown that inhibition of Nrf2 expression by RNAi suppressed tumor growth and induced sensitivity to platinum-based chemotherapy drugs (8, 20). We hypothesized that neoadjuvant platinum-based chemotherapy would lead to increased nuclear Nrf2 expression in NSCLC tumors; however, we found similar nuclear Nrf2 expression in both chemotherapy-naïve and chemotherapy-treated tumors, including those with pathologic characteristics that are associated with no response to neoadjuvant chemotherapy.

We observed a trend towards an association between nuclear Nrf2 expression and worse RFS and OS in patients with squamous cell carcinomas who had undergone surgery and received adjuvant platinum-based chemotherapy. Interestingly, this phenomenon was not observed in adenocarcinoma patients. These findings suggest that, as has been demonstrated in *in vitro* and *in vivo* studies using NSCLC cell lines (8, 20, 22, 34), nuclear Nrf2 expression in malignant lung cancer cells may play a role in chemotherapy resistance in squamous cell carcinoma subtype. However, these observations need to be studied further in a larger number of cases as part of prospectively conducted clinical trials. Importantly, the role of Nrf2 expression as a potential predictive marker associated to resistance to platinum-based chemotherapy needs to be addressed

in NSCLC patients with advanced (metastatic) tumors in which a more direct correlation between Nrf2 expression and response to chemotherapy can be established.

In summary, increased Nrf2 expression and decreased expression of Keap1 are common abnormalities in surgically resected NSCLCs and are associated with clinical outcome. In our study, abnormal expression of Nrf2 and Keap1 proteins was more common than that of the corresponding gene mutations, suggesting that other mechanisms are involved in the activation of *NFE2L2* and inactivation of *KEAP1*. Nrf2 expression may play a role in response to adjuvant platinum-based chemotherapy in patients with squamous cell carcinoma. Identifying patients with abnormal Nrf2 expression may be important for selection for chemotherapy in NSCLC, and our data suggest that Nrf2 expression could be added to the list of potential molecular markers to be tested to personalize treatment of NSCLC when platinum-based chemotherapeutic agents are used.

REFERENCES

1. Jemal A, Siegel R, Ward E, *et al.* Cancer statistics, 2008. *CA Cancer J Clin* 2008;58: 71-96.
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med* 2008;359: 1367-80.
3. Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon JP, Vansteenkiste J. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 2004;350: 351-60.
4. Bordoni R. Consensus conference: multimodality management of early- and intermediate-stage non-small cell lung cancer. *Oncologist* 2008;13: 945-53.
5. Singh A, Misra V, Thimmulappa RK, *et al.* Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer. *PLoS Med* 2006;3: e420.
6. Colburn NH, Kensler TW. Targeting transcription factors for cancer prevention--the case of Nrf2. *Cancer Prev Res (Phila Pa)* 2008;1: 153-5.
7. Lau A, Villeneuve NF, Sun Z, Wong PK, Zhang DD. Dual roles of Nrf2 in cancer. *Pharmacol Res* 2008;58: 262-70.
8. Singh A, Boldin-Adamsky S, Thimmulappa RK, *et al.* RNAi-mediated silencing of nuclear factor erythroid-2-related factor 2 gene expression in non-small cell lung cancer inhibits tumor growth and increases efficacy of chemotherapy. *Cancer Res* 2008;68: 7975-84.
9. Hayes JD, McMahon M. NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. *Trends Biochem Sci* 2009;34: 176-88.
10. Thimmulappa RK, Mai KH, Srisuma S, Kensler TW, Yamamoto M, Biswal S. Identification of Nrf2-regulated genes induced by the chemopreventive agent sulforaphane by oligonucleotide microarray. *Cancer Res* 2002;62: 5196-203.

11. Wang R, An J, Ji F, Jiao H, Sun H, Zhou D. Hypermethylation of the Keap1 gene in human lung cancer cell lines and lung cancer tissues. *Biochem Biophys Res Commun* 2008;373: 151-4.
12. Shibata T, Ohta T, Tong KI, *et al.* Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. *Proc Natl Acad Sci U S A* 2008;105: 13568-73.
13. Shibata T, Kokubu A, Gotoh M, *et al.* Genetic alteration of Keap1 confers constitutive Nrf2 activation and resistance to chemotherapy in gallbladder cancer. *Gastroenterology* 2008;135: 1358-68, 68 e1-4.
14. Kim YR, Oh JE, Kim MS, *et al.* Oncogenic NRF2 mutations in squamous cell carcinomas of oesophagus and skin. *J Pathol* 2009;220: 446-51.
15. Eberhard DA, Johnson BE, Amler LC, *et al.* Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol* 2005;23: 5900-9.
16. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. *Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart*. Lyon: International Agency for Research on Cancer (IARC); 2004. p. 9-124.
17. Sauter G, Simon R, Hillan K. Tissue microarrays in drug discovery. *Nat Rev Drug Discov* 2003;2: 962-72.

18. Shigematsu H, Lin L, Takahashi T, *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005;97: 339-46.
19. Tang X, Shigematsu H, Bekele BN, *et al.* EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res* 2005;65: 7568-72.
20. Cho JM, Manandhar S, Lee HR, Park HM, Kwak MK. Role of the Nrf2-antioxidant system in cytotoxicity mediated by anticancer cisplatin: implication to cancer cell resistance. *Cancer Lett* 2008;260: 96-108.
21. Ohta T, Iijima K, Miyamoto M, *et al.* Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. *Cancer Res* 2008;68: 1303-9.
22. Kim HR, Kim S, Kim EJ, *et al.* Suppression of Nrf2-driven heme oxygenase-1 enhances the chemosensitivity of lung cancer A549 cells toward cisplatin. *Lung Cancer* 2008;60: 47-56.
23. Padmanabhan B, Tong KI, Ohta T, *et al.* Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. *Mol Cell* 2006;21: 689-700.
24. Kim JH, Bogner PN, Ramnath N, Park Y, Yu J, Park YM. Elevated peroxiredoxin 1, but not NF-E2-related factor 2, is an independent prognostic factor for disease recurrence and reduced survival in stage I non-small cell lung cancer. *Clin Cancer Res* 2007;13: 3875-82.
25. Stacy DR, Ely K, Massion PP, *et al.* Increased expression of nuclear factor E2 p45-related factor 2 (NRF2) in head and neck squamous cell carcinomas. *Head Neck* 2006;28: 813-8.
26. Itoh K, Wakabayashi N, Katoh Y, *et al.* Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev* 1999;13: 76-86.

27. Dhakshinamoorthy S, Jaiswal AK. Functional characterization and role of INrf2 in antioxidant response element-mediated expression and antioxidant induction of NAD(P)H:quinone oxidoreductase1 gene. *Oncogene* 2001;20: 3906-17.
28. Lo SC, Li X, Henzl MT, Beamer LJ, Hannink M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *Embo J* 2006;25: 3605-17.
29. Li W, Kong AN. Molecular mechanisms of Nrf2-mediated antioxidant response. *Mol Carcinog* 2009;48: 91-104.
30. Karapetian RN, Evstafieva AG, Abaeva IS, *et al.* Nuclear oncoprotein prothymosin alpha is a partner of Keap1: implications for expression of oxidative stress-protecting genes. *Mol Cell Biol* 2005;25: 1089-99.
31. Lo SC, Hannink M. PGAM5, a Bcl-XL-interacting protein, is a novel substrate for the redox-regulated Keap1-dependent ubiquitin ligase complex. *J Biol Chem* 2006;281: 37893-903.
32. Strachan GD, Ostrow LA, Jordan-Sciutto KL. Expression of the fetal Alz-50 clone 1 protein induces apoptotic cell death. *Biochem Biophys Res Commun* 2005;336: 490-5.
33. Ronkainen H, Vaarala MH, Kauppila S, *et al.* Increased BTB-Kelch type substrate adaptor protein immunoreactivity associates with advanced stage and poor differentiation in renal cell carcinoma. *Oncol Rep* 2009;21: 1519-23.
34. Jozkowicz A, Was H, Dulak J. Heme oxygenase-1 in tumors: is it a false friend? *Antioxid Redox Signal* 2007;9: 2099-117.

FIGURE LEGENDS

Figure 1. Microphotographs showing positive (*blue arrows*) and negative nuclear Nrf2 expression; and absent, low and high cytoplasmic Keap1 expression in NSCLC tumor tissue specimens: squamous cell carcinoma (*A, C to E*), and adenocarcinoma (*B and F*) (magnification, x100, except *C* x200).

Figure 2. *A*, Five-year OS and *B*, RFS rated by nuclear Nrf2 protein expression in all patients with NSCLC. *C*, Five-year OS and *D*, RFS rates by cytoplasmic Keap1 protein expression in patients with squamous cell carcinoma. *E*, Five-year OS and *F*, RFS rated by nuclear Nrf2 and low cytoplasmic Keap1 proteins expression in all patients with NSCLC. (*E*, events; and *N*, total number of cases).

Supplementary Figure 1. Nrf2 and Keap1 protein expression analysis by Western blot (A) and immunohistochemical analysis in NSCLC cell lines and immortalized bronchial epithelial BEAS2B cells (B). A: On Western blot analysis, Nrf2-positive NSCLC cell lines had higher expression in the nucleus (N) than in the cytoplasm (C). Poly (ADP-ribose) polymerase expression (PARP) indicates nuclear protein lysates. B: Microphotographs (magnification, x400) showing nuclear Nrf2 expression (*blue arrows*) in 2 cell lines (H460 and A549); the other 2 cell lines (H1933 and BEAS2B) demonstrated cytoplasmic expression only. Microphotographs (magnification, x400) showing cytoplasmic Keap1 (*red arrows*) expression in 2 cell lines (H1933 and BEAS2B); the other 2 cell lines (H460 and A549) showed low cytoplasmic Keap1 expression.

Supplementary Figure 2. A, Five-year OS and B, RFS rated by nuclear Nrf2 protein expression in adenocarcinoma patients. C, Five-year OS and D, RFS rates by nuclear Nrf2 protein expression in patients with squamous cell carcinoma. (*E*, events; and *N*, total number of cases).

Supplementary Figure 3. A, Five-year OS and B, RFS rated by cytoplasmic Keap1 protein expression in all NSCLC patients. C, Five-year OS and D, RFS rates by cytoplasmic Keap1 protein expression in patients with adenocarcinoma. (*E*, events; and *N*, total number of cases).

Supplementary Figure 4. A, Five-year RFS rated by nuclear Nrf2 protein expression squamous cell carcinoma patients treated with adjuvant therapy. B, F Five-year RFS rated by nuclear Nrf2 protein expression squamous cell carcinoma patients not treated with adjuvant therapy. (*E*, events; and *N*, total number of cases).

Fig. 1

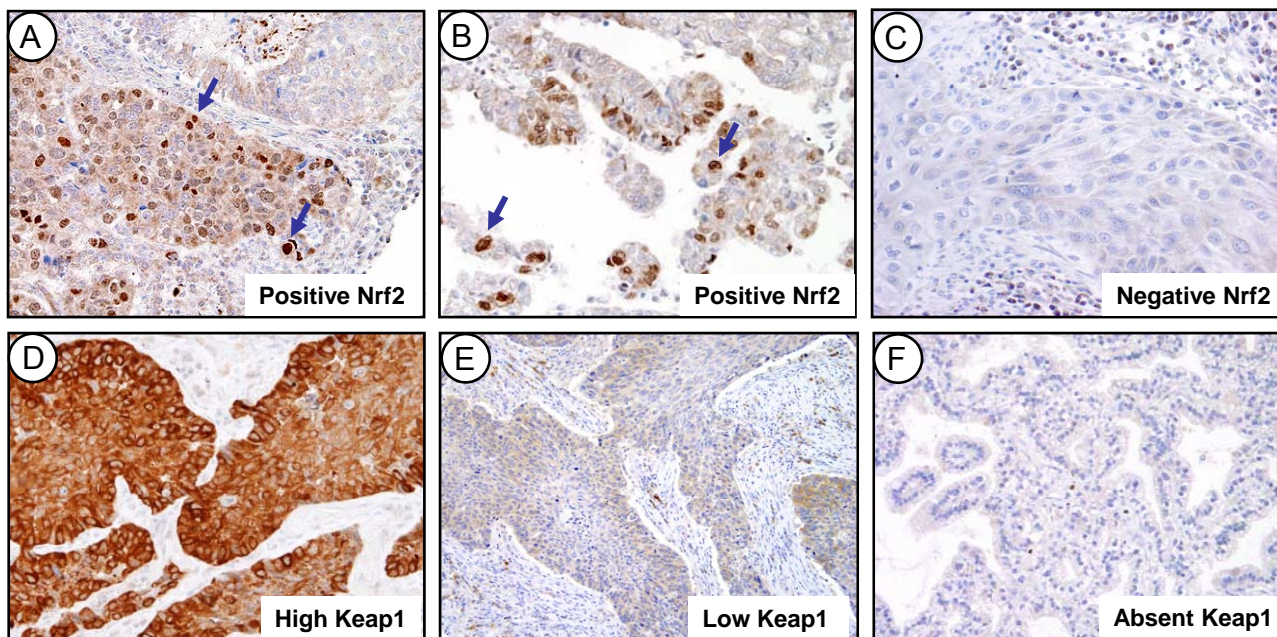


Fig. 2

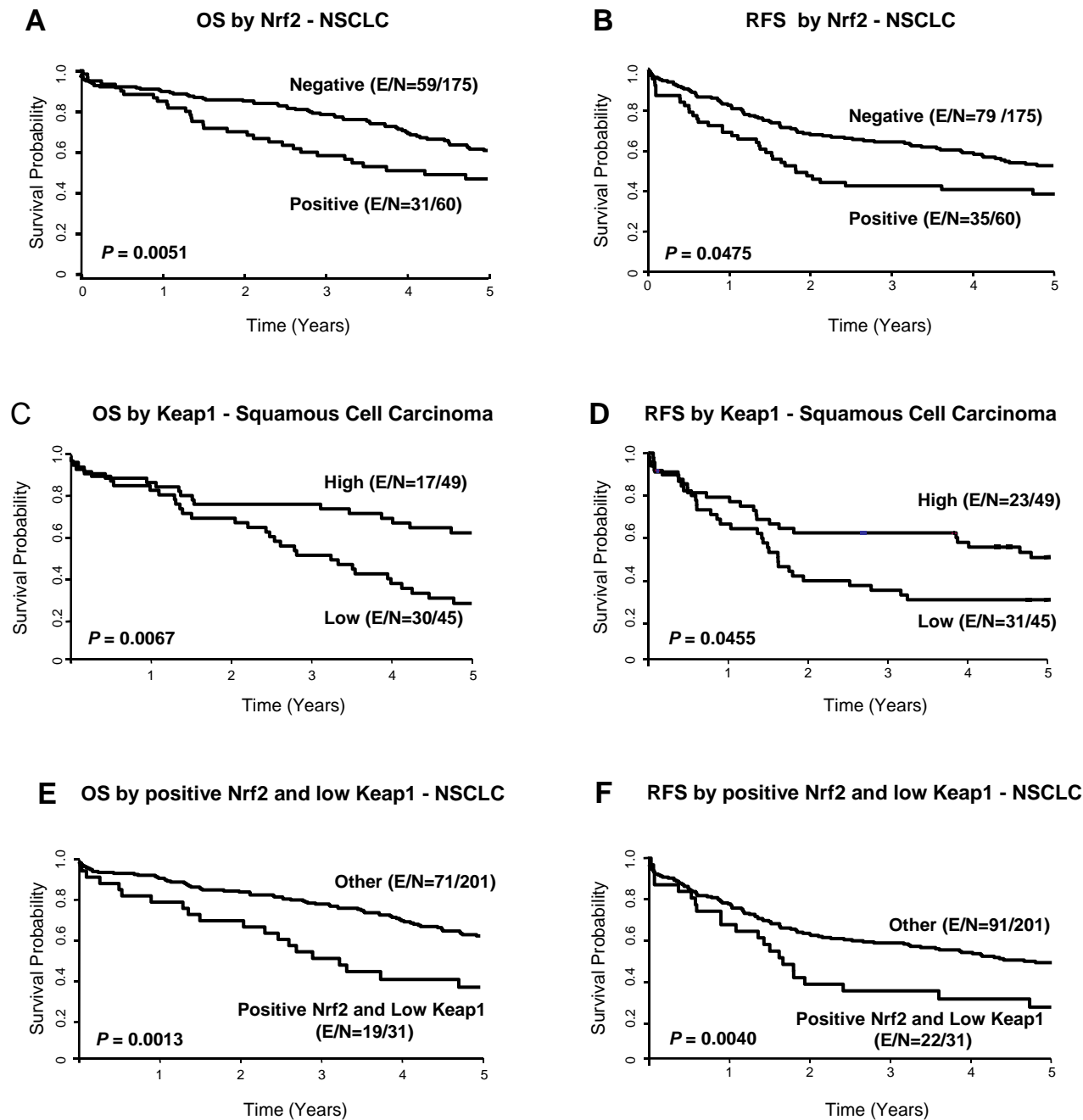


Table 1. Clinicopathologic features of NSCLCs evaluated for Nrf2 and Keap1 expression in TMA

Characteristic [#]		Total (n=304)	Adenocarcinoma (n=190)	Squamous cell carcinoma (n=114)
Age Mean (years)		66	65	68
Sex				
	Female	157	113	44
	Male	147	77	70
Pathologic stage				
	I	191	127	64
	II	58	24	34
	III	46	32	14
	IV	9	7	2
Smoking history [†]				
	Current/Former	253	144	109
	Never	50	46	4

[†]Smoking history was not available for 1 patient.

[#] ECOG (Eastern Cooperative Oncologic Group) performance status of the patients was 0 and 1

Table 2. Nuclear Nrf2 expression according to NSCLC histology type and chemotherapy treatment

Sample Type	Adenocarcinoma Positive Nrf2[†]/ Total (%)	Squamous Cell Carcinoma Positive Nrf2[†]/ Total (%)	NSCLC Positive Nrf2[†]/ Total (%)	<i>P</i> Value[#]
Tissue Microarray	34/188 (18%)	43/112 (38%)	77/300 (26%)	<0.0001
Whole Sections				
Neoadjuvant Treated	4/20 (20%)	3/6 (50%)	7/26 (27%)	0.2929
Adjuvant Treated	13/35 (37%)	14/28 (50%)	27/63 (43%)	0.3055
No Adjuvant Treated	10/27 (37%)	10/32 (31%)	20/59 (34%)	0.6399

[†] Positive Nrf2, score >0.[#] *P* value comparison of the frequencies of marker expression

Table 3. Multivariate Cox model of 5-year overall survival (OS) and recurrence-free survival (RFS) in patients with NSCLC and squamous cell carcinoma.

Variable	HR (95% CI)	P value
OS, Nrf2 expression, NSCLC		
Age at surgery (per 1-year increase)	1.067 (1.043-1.092)	<0.0001
Stage (II vs. I)	2.096 (1.332-3.299)	0.0015
Smoking (current and former vs. never)	2.476 (1.069-5.736)	0.0331
Nuclear Nrf2 (positive vs. negative)	1.747 (1.120-2.726)	0.0007
OS, Keap1 expression, squamous cell carcinoma		
Age at surgery (per 1-year increase)	1.054 (1.020-1.090)	0.0020
Stage (II vs. I)	1.876 (1.033-3.409)	0.0389
Cytoplasmic Keap1 (score <150 vs. ≥150)	2.087 (1.134-3.841)	0.0181
OS, Nrf2 and Keap1 expression, NSCLC		
Age at surgery (per 1-year increase)	1.063 (1.039-1.087)	<0.0001
Stage (II vs. I)	2.093 (1.332-3.290)	0.0014
Smoking (current and former vs. never)	2.639 (1.134-6.142)	0.0244
Nrf2/Keap1 status (positive Nrf2 and Keap1 <150 vs. other)	1.966 (1.167-3.313)	0.0111
RFS, Nrf2 and Keap1 expression, NSCLC		
Age at surgery (per 1-year increase)	1.043 (1.023-1.064)	<0.0001
Stage (II vs. I)	1.962 (1.296-2.970)	0.0014
Smoking (current and former vs. never)	2.115 (1.064-4.203)	0.0325
Nrf2/Keap1 status (positive Nrf2 and Keap1 <150 vs. other)	1.688 (1.045-2.727)	0.0325



Knockdown of *ZEB1*, a master epithelial-to-mesenchymal transition (EMT) gene, suppresses anchorage-independent cell growth of lung cancer cells

Yoshihiro Takeyama^a, Mitsuo Sato^{a,*}, Mihoko Horio^a, Tetsunari Hase^a, Kenya Yoshida^a, Toshihiko Yokoyama^c, Harunori Nakashima^a, Naozumi Hashimoto^a, Yoshitaka Sekido^{b,d}, Adi F. Gazdar^e, John D. Minna^e, Masashi Kondo^a, Yoshinori Hasegawa^a

^a Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

^b Department of Cancer Genetics, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan

^c Department of Respiratory Medicine, Japan Red Cross Nagoya First Hospital, 3-35 Michishita-cho, Nakamura-ku, Nagoya, Aichi 453-8511, Japan

^d Division of Molecular Oncology, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan

^e Hamon Center for Therapeutic Oncology Research and the Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center at Dallas, 6000 Harry Hines, Dallas, TX 75390-8593, USA

ARTICLE INFO

Article history:

Received 5 January 2010

Received in revised form 7 April 2010

Accepted 7 April 2010

Available online xxxx

Keywords:

Lung cancer

Epidermal growth factor receptor

Anchorage-independent growth

EMT

MicroRNA

RNA interference

ABSTRACT

We found that among four master epithelial-to-mesenchymal transition (EMT)-inducing genes (*ZEB1*, *SIP1*, *Snail*, and *Slug*) *ZEB1* expression was most significantly correlated with the mesenchymal phenotype (high *Vimentin* and low *E-cadherin* expression) in non-small cell lung cancer (NSCLC) cell lines and tumors. Furthermore, *ZEB1* knockdown with RNA interference in three NSCLC cell lines with high *ZEB1* expression suppressed to varying degrees mass culture growth and liquid colony formation but in all cases dramatically suppressed soft agar colony formation. In addition, *ZEB1* knockdown induced apoptosis in one of the three lines, indicating that the growth inhibitory effects of *ZEB1* knockdown occurs in part through the activation of the apoptosis pathway. These results suggest that inhibiting *ZEB1* function may be an attractive target for NSCLC therapeutic development.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Epithelial-to-mesenchymal transition (EMT) is an embryonic developmental program involving changes in cell morphology and expression of EMT-associated genes [1,2]. EMT also occurs during the progression of several types of human cancer and confers motility and invasiveness on cancer cells, leading them to acquire ability to metastasize to distant sites. Genetics of early development discovered a number of EMT-inducing genes encoding transcription factors capable of inducing EMT when ectopically expressed in epithelial cells. Several EMT-inducing

genes that have essential roles in EMT are called master EMT genes, including *Twist*, *ZEB1*, *SIP1*, *Snail*, *Slug*, and *Gooseoid* [1,2]. These genes function as transcriptional repressors of the cell–cell adhesion glycoprotein, *E-cadherin* whose functional loss is one of the hallmarks of EMT [3]. Among these master EMT genes, *Snail* was shown to repress *E-cadherin* and to induce EMT in cancer cells [4,5], while *Twist* was demonstrated to promote breast cancer metastasis [6]. Increased expression of *Twist* and *Snail* have been shown in hepatocellular, breast, colorectal, and gastric cancers, often correlating with poor prognosis [7–10].

Lung cancer is the leading cause of cancer deaths, killing over 1 million people every year worldwide [11]. It develops through a multi-step process involving accumulation

* Corresponding author. Tel.: +81 52 744 2167; fax: +81 52 744 2176.
E-mail address: msato@med.nagoya-u.ac.jp (M. Sato).

of multiple genetic and epigenetic changes that confer growth advantages on normal lung epithelial cells, leading them to transform to clinically evident lung cancer cells [12,13]. Analyzing a large number of lung tumor specimens, Prudkin et al. showed that the majority of primary lung cancers and even premalignant lesions have the mesenchymal phenotype as characterized by down-regulation of E-cadherin and up-regulation of Vimentin [14]. Although several master EMT genes have been shown to directly contribute to tumor progression in breast, colon, and pancreatic cancers, very little is known about functional roles of master EMT genes in lung cancer progression.

In addition to *Snail* and *Twist*, recently *ZEB1* has emerged as a key player in cancer progression. *ZEB1* promotes tumor metastasis in colon and breast cancer [15], is associated with resistance to conventional chemotherapy in pancreatic cancer [16,17], and potentially has a predominant role in inducing EMT in NSCLC. First, among several transcription factors including *ZEB1*, *Snail*, and β -catenin, *ZEB1* protein expression showed the most significant inverse correlation with E-cadherin in NSCLC and mesothelioma cell lines [18]. Second, prostaglandin E2 was shown to exert its ability to suppress E-cadherin through inducing *ZEB1* and *Snail* in lung cancer cell lines [19]. Third, *ZEB1* has been shown to suppress the Semaphorin 3F tumor suppressor gene in lung cancer cells [20]. These suggest relevant roles of *ZEB1* as an EMT-inducer as well as an oncogene in lung cancer.

Thus, we performed this study aiming to evaluate the association between *ZEB1* expression and the mesenchymal phenotype in lung cancer, and to test the effects of *ZEB1* knockdown with RNA interference on the growth of lung cancer cells. We found that *ZEB1* expression significantly correlates with increased *Vimentin* and decreased *E-cadherin* expression in lung cancer, while knockdown of *ZEB1* resulted in dramatic growth inhibition in lung cancer cell lines. These results suggest that *ZEB1* is a promising therapeutic target for lung cancer.

2. Materials and methods

2.1. Cell lines and primary tumor tissues

NSCLC cell lines used in this study were purchased from American Type Culture Collection or obtained from the Hamon Center collection (University of Texas Southwestern Medical Center). These cells include PC9, A549, NCI-H157, NCI-H460, NCI-H820, NCI-H838, NCI-H1155, NCI-H1299, NCI-H1666, NCI-H1650, NCI-H1975, NCI-H3255, HCC44, HCC827, HCC2279, HCC2935, HCC4006, and HCC4011 (cells with mutation in epidermal growth factor receptor (EGFR) gene are underlined) [21]. A mesothelioma cell line, ACC-MESO-1, which was used as positive control for western blot of cleaved caspase-3, was established by ourselves [22]. Cells were cultured with RPMI 1640 (Sigma–Aldrich Corp, MO, USA) supplemented with 10% fetal bovine serum. Surgically resected 32 primary tumor specimens (19 adenocarcinomas and 13 squamous cell carcinomas) were obtained from patients at the Nagoya University

Hospital, Nagoya First Japan Red Cross Hospital, Nagoya Second Japan Red Cross Hospital, Kasugai Municipal Hospital and Chukyo Hospital in Nagoya, Japan. Before tissue samples were collected ethical approval of the each institute and fully informed written consents from all patients were obtained. We previously analyzed *EGFR* mutation status of these samples and used the data of the analysis for the present study [23].

2.2. RNA isolation and quantitative real-time PCR analysis

For mRNA analysis, 5 μ g of total RNA isolated using Trizol (Invitrogen Corp., CA, USA) were reverse transcribed with Super script III First-Strand Synthesis System using Random primer system (Invitrogen Corp.). Quantitative real-time PCR (qRT-PCR) analysis of *E-cadherin*, *Vimentin*, *ZEB1*, *SIP1*, *Snail*, and *Slug*, was performed as described previously using the standard Taqman assay-on-demand PCR protocol in a reaction volume of 20 μ L, including 50 ng cDNA [24]. We used the comparative C_t method to compute relative expression values. For microRNA analysis, 10 ng of total RNA isolated using *mirVana* miRNA Isolation Kit (Applied Biosystems, CA, USA) were reverse transcribed with TaqMan MicroRNA Reverse Transcription Kit using a primer set specific for each of microRNAs (miR-200a, miR-200b, miR-200c, and miR-205) studied (Applied Biosystems). qRT-PCR analysis of microRNA was done as described above. We used *GAPDH* (Applied Biosystems assay-on-demand) for mRNA analysis and U6 small nuclear (sn) RNA for microRNA analysis as internal controls.

2.3. Western blot analysis

Western blot analysis was done as described previously using whole cell lysates [24]. Primary antibodies used were mouse monoclonal anti-E-cadherin, anti-Vimentin (BD Bioscience, NJ, USA), goat polyclonal anti-ZEB1 (Santa Cruz biotech., CA, USA), and rabbit polyclonal anti-cleaved caspase-3 (Cell Signaling Tech., MA, USA). Actin protein levels were used as a control for adequacy of equal protein loading. Anti-rabbit, anti-mouse (GE healthcare, Buckinghamshire, England), or anti-goat antibody (R&D Systems, MN, USA) was used at 1:2000 dilution as a secondary antibody.

2.4. Transfection of short interfering RNA

The 4.5×10^5 of H1299 and H157 or 9.0×10^5 of H460 cells were plated in 6-well plates. Next day, cells were transiently transfected with either 10 nM predesigned short interfering RNA (siRNA) (Stealth Select RNAi) targeting *ZEB1* or control siRNA purchased from Invitrogen using Lipofectamine RNAiMAX (Invitrogen Corp.) according to the manufacturer's protocol. After 48 h, the transfected cells were harvested for further analyses or plated for cell growth assays.

2.5. Cell growth assays

Colorimetric proliferation assay was performed using WST-1 assay kit (Roche, Basel, Switzerland) according to

manufacturer's instruction. Liquid and soft agar colony formation assays were done as described previously [24].

2.6. Cell cycle analysis

Cells were harvested 48 h after the transfection of siRNA oligos. Cells were, fixed, treated with RNase A, stained with propidium iodide using BD Cycletest Plus Reagent Kit (BD Bioscience) according to the instructions of the manufacture, and analyzed by flow cytometry for DNA

synthesis and cell cycle status [FACSCalibur instrument, (Becton Dickinson) with BD CellQuest™Pro Ver.5.2.1 (BD Bioscience)].

2.7. Senescence associated β -galactosidase staining

Cells were stained with β -galactosidase using Senescence β -Galactosidase Staining Kit (Cell Signaling Tech.), and cells stained blue were counted under a microscope (200 \times total magnification).

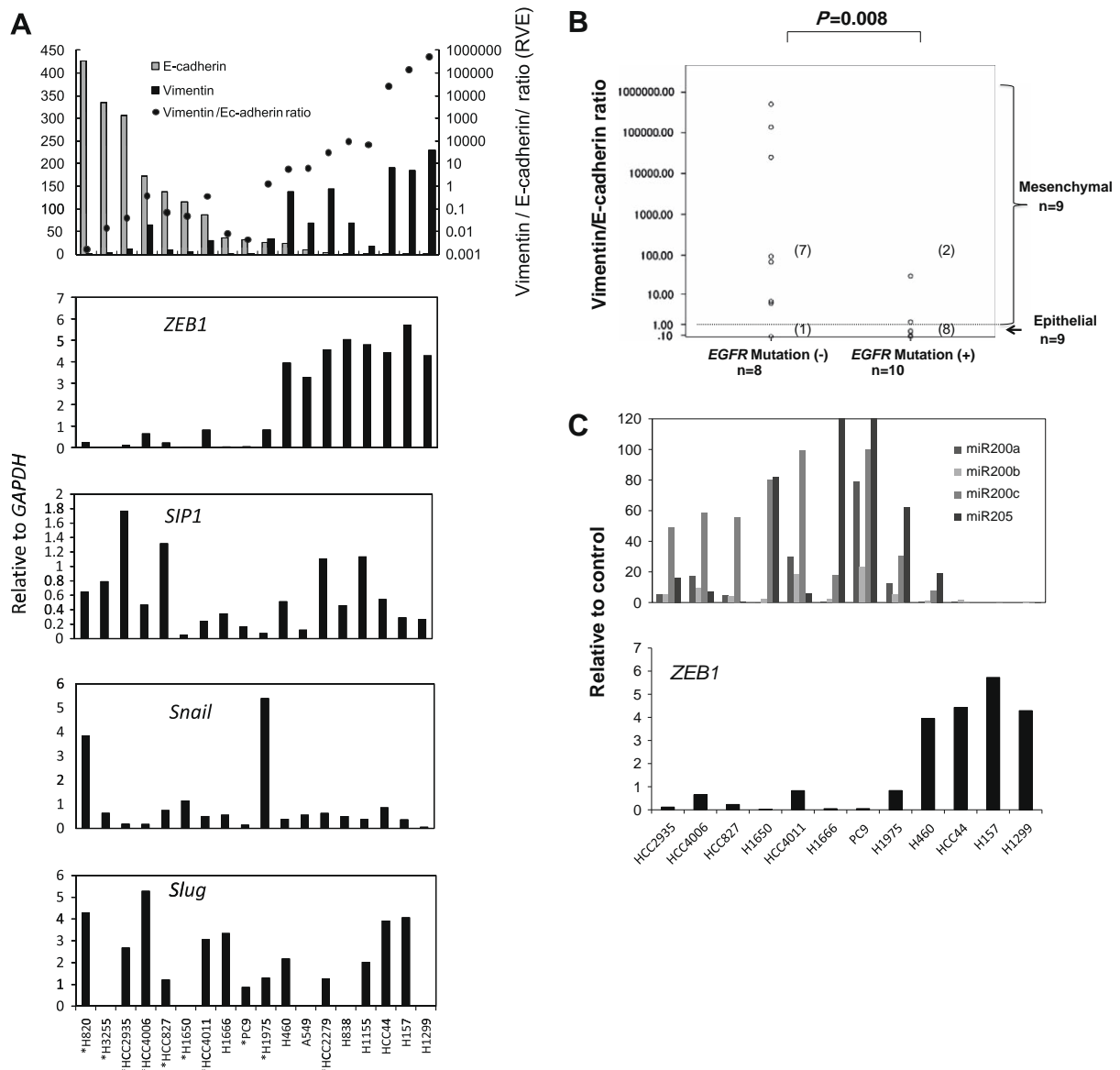


Fig. 1. *ZEB1* expression significantly correlates with *E-cadherin*, *Vimentin* and two microRNAs, miR-200c and miR-205, known *ZEB1* repressor in lung cancer cell lines. (A) Quantitative real-time PCR (qRT-PCR) analysis of *E-cadherin*, *Vimentin* and four master EMT genes, *ZEB1*, *SIP1*, *Snail*, and *Slug* in 18 lung cancer cell lines. The cell lines are aligned by expression levels of *E-cadherin* from high (left) to low (right). The results are average of two independent PCR experiments done in duplicated reactions. Statistic values of correlations between these genes are shown in Table 1. *Cells with mutation in the epidermal growth factor receptor (*EGFR*) gene. (B) The association between the ratio of *Vimentin* to *E-cadherin* (RVE) and mutation status of *EGFR* in lung cancer cell lines. (C) qRT-PCR analysis of four microRNAs known as repressors of *ZEB1*. The expression levels of miR-200c and miR-205 were significantly negatively associated with *ZEB1* expression (Spearman's correlation coefficients: -0.71 , $P = 0.01$ for miR-200c and -0.79 , $P = 0.002$ for miR-205).

2.8. Statistics

SPSS ver.17 software was used for all statistic analyses in this study. Spearman's correlation coefficients with associated *P* values were calculated between mRNA expression of *Vimentin*, *E-cadherin*, and four EMT-inducing genes, microRNA expression of the four microRNAs, and the ratio of *Vimentin* to *E-cadherin* expression (RVE). Mann–Whitney *U* test was used for analyzing difference between two groups.

3. Results

3.1. Lung cancer cell lines can be divided into epithelial and mesenchymal phenotypes based on the expression status of *E-cadherin* and *Vimentin*

We first examined whether lung cancer cell lines can be classified into mesenchymal and epithelial phenotypes based on expression status of *E-cadherin* and *Vimentin*, which are markers for epithelial and mesenchymal phenotypes, respectively. qRT-PCR analysis of *E-cadherin* and *Vimentin* in 18 non-small cell lung cancer cell (NSCLC) lines including 10 with mutations in the epidermal growth factor receptor (*EGFR*) gene showed that most predominantly expressed either *E-cadherin* or *Vimentin* (Fig. 1A). We classified NSCLCs as either “epithelial” (high *E-cadherin*/low *Vimentin*) or “mesenchymal” (high *Vimentin*/low *E-cadherin*) according to the expression status of *E-cadherin* and *Vimentin*. We quantified the ratio of *Vimentin* to *E-cadherin* expression (RVE) as an index that represents the degree of mesenchymal phenotype, with NSCLCs showing RVE ≥ 1.0 classified as mesenchymal phenotype ($n = 9$ NSCLCs) and those with RVE < 1.0 as epithelial phenotype ($n = 9$ NSCLCs). Notably, RVEs of *EGFR* wild-type NSCLCs were significantly higher than those of *EGFR* mutant NSCLCs (Fig. 1B) (Median; 80.0 vs. 0.06, $P = 0.008$, Mann–Whitney *U* test). All but one *EGFR* wild-type NSCLC lines were “mesenchymal” while 8 of 10 *EGFR* mutant NSCLCs were “epithelial”, suggesting that *EGFR* mutation is associated with epithelial characteristics (Fig. 1B).

3.2. Among four master EMT genes only *ZEB1* expression was significantly correlated with both *Vimentin* and *E-cadherin* expression in lung cancer cell lines

To identify master EMT genes whose expressions are significantly associated with the mesenchymal phenotype in lung cancer cell lines, we analyzed the expression levels of four master EMT genes (*ZEB1*, *SIP1*, *Snail*, and *Slug*) (Fig. 1A, Table 1). *ZEB1* expression was inversely correlated with *E-cadherin* expression (Spearman's correlation coefficient = -0.82 , $P < 0.001$), and positively correlated with *Vimentin* expression (Spearman's correlation coefficient = 0.80 , $P < 0.001$) (Table 1), resulting in

highly significant association between *ZEB1* expression and RVE (Spearman's correlation coefficient = 0.88 , $P < 0.001$) (Table 1). By contrast, *SIP1*, *Snail*, and *Slug* expression were not correlated with *E-cadherin* expression, *Vimentin* expression, or RVE (Table 1). These results suggested that *ZEB1* may induce EMT in lung cancer.

3.3. The expression levels of four microRNAs known as repressors of *ZEB1* were negatively correlated with *ZEB1* expression in NSCLCs

Recently, several groups have shown that *ZEB1* expression is down-regulated by three members of microRNA-200 family (miR-200a, miR-200b, miR-200c) and miR-205 in different types of cancer cells [25–28]. To confirm these findings in a panel of lung cancer cell lines, we performed qRT-PCR of the four microRNAs in 12 NSCLC cell lines and correlated their expression with *ZEB1* expression (Fig. 1C). We found that the expression levels of miR-200c and miR-205 were negatively correlated with *ZEB1* expression (Spearman's correlation coefficients: -0.71 , $P = 0.01$ for miR-200c, -0.79 , $P = 0.002$ for miR-205), while miR-200a and miR-200b were not correlated with *ZEB1* expression. These results suggest that miR-200c and miR-205 may have a major role in regulating *ZEB1* expression in lung cancer (Fig. 1C).

3.4. *ZEB1* and *Snail* expression were significantly correlated with RVE in primary lung tumor tissues

Next, we analyzed the expression of *E-cadherin*, *Vimentin*, and the four master EMT genes in 32 NSCLC tumor specimens and found that *ZEB1* expression was highly correlated with *Vimentin* expression (Spearman's correlation coefficient = 0.92 , $P < 0.001$) (Table 2). *ZEB1* (Spearman's correlation coefficient = 0.51 , $P = 0.003$), and *Snail* (Spearman's correlation coefficient = 0.45 , $P = 0.01$) expression were significantly correlated with RVE. However, we note that overall in the tumor specimens in contrast to the NSCLC lines, *E-cadherin* and *Vimentin* expression were positively correlated (Spearman's correlation coefficient = 0.63 , $P < 0.001$) (Table 2). This may be because the primary tumor specimens were macrodissected and thus included transcripts from both tumor and non-tumorous lung tissue while the NSCLC lines only reflected tumor expression. Probably because of this reason, unlike in the NSCLC lines, *EGFR* mutation status was not correlated with epithelial characteristics (high *E-cadherin* and low *Vimentin*) in the primary tumors (data not shown).

3.5. Knockdown of *ZEB1* induced significant suppression of anchorage-independent cell growth in lung cancer cell lines

The results presented above suggest that *ZEB1* plays a dominant role in maintaining the mesenchymal phenotype in NSCLCs. To test the therapeutic potential of *ZEB1* we performed RNA interference (RNAi)-mediated gene silencing against *ZEB1*. To minimize the possibility of “off target effect”, we used low dose (10 nM) Stealth Select RNAi (Invitrogen),

Table 1

Correlations between mRNA expression of master EMT genes, *E-cadherin*, and *Vimentin* in 18 non-small lung cancer cell lines.

	<i>E-cadherin</i>	<i>Vimentin</i>	RVE	<i>ZEB1</i>	<i>SIP1</i>	<i>Snail</i>
<i>E-cadherin</i>						
<i>Vimentin</i>	-0.74					
	<0.001					
RVE	-0.88	0.92				
	<0.001	<0.001				
<i>ZEB1</i>	-0.82	0.80	0.88			
	<0.001	<0.001	<0.001			
<i>SIP1</i>	0.22	-0.05	-0.06	-0.06		
	0.38	0.85	0.81	0.80		
<i>Snail</i>	0.30	-0.22	-0.26	-0.23	-0.03	
	0.23	0.38	0.30	0.36	0.92	
<i>Slug</i>	0.09	0.07	-0.04	0.20	0.24	-0.06
	0.72	0.79	0.88	0.46	0.34	0.82

Spearman's correlation coefficients (upper row) and statistic values (lower row) are shown. Statistically significant correlations ($P < 0.01$) are in bold. RVE; the ratio of *Vimentin* to *E-cadherin*.

Table 2

Correlations between mRNA expression of master EMT genes, *E-cadherin*, and *Vimentin* in 32 non-small lung cancer tumor tissues.

	<i>E-cadherin</i>	<i>Vimentin</i>	RVE	<i>ZEB1</i>	<i>SIP1</i>	<i>Snail</i>
<i>E-cadherin</i>						
<i>Vimentin</i>	0.63					
	<0.001					
RVE	-0.26	0.53				
	0.15	0.002				
<i>ZEB1</i>	0.57	0.92	0.51			
	<0.001	<0.001	0.003			
<i>SIP1</i>	0.81	0.73	0.09	0.64		
	<0.001	<0.001	0.61	<0.001		
<i>Snail</i>	0.32	0.71	0.45	0.69	0.49	
	0.07	<0.001	0.01	<0.001	0.004	
<i>Slug</i>	0.29	0.39	0.31	0.41	0.43	0.26
	0.11	0.03	0.08	0.02	0.01	0.15

Spearman's correlation coefficients (upper row) and statistic values (lower row) are shown. Statistically significant correlations ($P < 0.01$) are in bold. RVE; the ratio of *Vimentin* to *E-cadherin*.

which includes three short interfering RNA (siRNA) oligos with non-overlapping sequences targeting *ZEB1*. We studied three NSCLC cell lines, H1299, H460, and H157, all of which express high levels of *ZEB1* mRNA (Fig. 1A). These cell lines were transiently transfected with each of the three *ZEB1* siRNAs or control oligos and harvested for analyses 48 h after transfection. Western blot of *ZEB1* showed that with all three oligos clear suppression of *ZEB1* protein was obtained in all cell lines (Fig. 2A). Western blot of E-cadherin, which is a direct negative transcription target of *ZEB1*, showed reexpression of E-cadherin protein in H460 but not in H1299 or H157 after the transfection (Fig. 2A). This was likely to reflect the difference in basal expression of E-cadherin protein between the cells; H460 but not H1299 or H157 expressed detectable levels of E-cadherin protein (Fig. 2A). However, we also considered the possibility that *ZEB1* knockdown did not upregulate *E-cadherin* mRNA in H1299 or H157 and thus performed qRT-PCR analysis of *E-cadherin*. The analysis revealed that after the *ZEB1* knockdown transfection *E-cadherin* mRNA significantly increased in H157 and H460 but remained unchanged in H1299, indicating that in H1299 *E-cadherin* mRNA was not upregulated by *ZEB1* knockdown

(Fig. 2B). Small or no changes in Vimentin protein expression were seen after the transfection of *ZEB1* knockdown oligos (Fig. 2A). Finally, we did not see significant morphologic changes suggestive of the cells undergoing EMT (data not shown).

We found *ZEB1* knockdown to have modest effects suppressing mass culture and liquid colony formation growth but much greater effects suppressing anchorage-independent growth in soft agar (Fig. 2C–E). This result indicates that the ability of these cells to grow in soft agar is highly dependent on *ZEB1* expression.

3.6. Growth inhibitory effect of *ZEB1* knockdown in lung cancer was caused in part by apoptosis but not by induction of senescence

The NSCLC cell lines varied in the induction of apoptosis following *ZEB1* knockdown with H460 cells showing the largest amount of *ZEB1* knockdown-induced apoptosis (Fig. 3A and B). One study has showed that mouse embryonic fibroblasts derived from *Zeb1* $-/-$ mice underwent

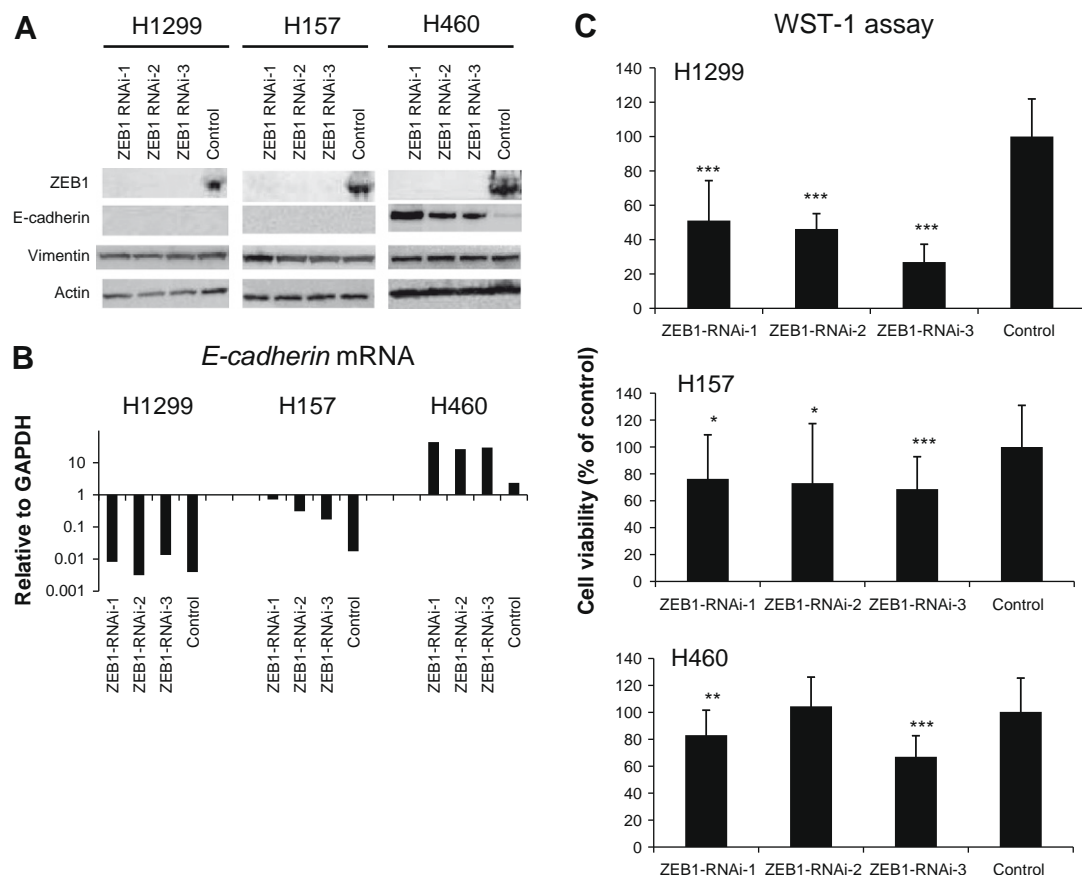


Fig. 2. RNAi-mediated knockdown of *ZEB1* suppresses cell proliferation, liquid colony formation (anchorage-dependent condition), and soft agar colony formation (anchorage-independent condition) in lung cancer cells. (A) Western blot of *ZEB1*, E-cadherin, and Vimentin for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. Actin was used as a loading control. The result is a representative result of two independent experiments. (B) qRT-PCR analysis of *E-cadherin* for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. The result is an average of three independent experiments done in duplicates. (C) Cell proliferation (WST-1) assay of lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. Cells were transfected with three different RNAi oligos targeting *ZEB1* or control oligos. Forty-eight hours after the transfection cells were counted and 1000 cells were plated in 96-well plates. Absorbance values were determined 96 h after the transfection. Results are from three independent experiments in eight replicates and shown as mean \pm SD. Values of cells transfected with control oligos are set as 100%. (D) Liquid colony formation assay for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. Forty-eight hours after the transfection cells were counted and 500 or 1000 cells were plated in 6-well plates in triplicates, cultured for two weeks, and stained with methylene blue. Colonies >2 mm were counted. Results are from three independent experiments and shown as mean \pm SD. Colony numbers of cells transfected with control oligos are set as 100%. (E) Soft agar colony formation assay for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. Forty-eight hours after the transfection cells were counted and 1000 cells were suspended in 0.37% SeaKem GTG Agarose (Lonza, Rockland, ME, USA) in triplicate 12-well plates. About 2 weeks later, colonies (>50 cells) were counted. Results are from three independent experiments and shown as mean \pm SD. Colony numbers of cells transfected with control oligos are set as 100%. In Fig. 2C–E, *, **, and *** indicate $P < 0.05$, $P < 0.01$, and $P < 0.001$ (Mann–Whitney U test), respectively.

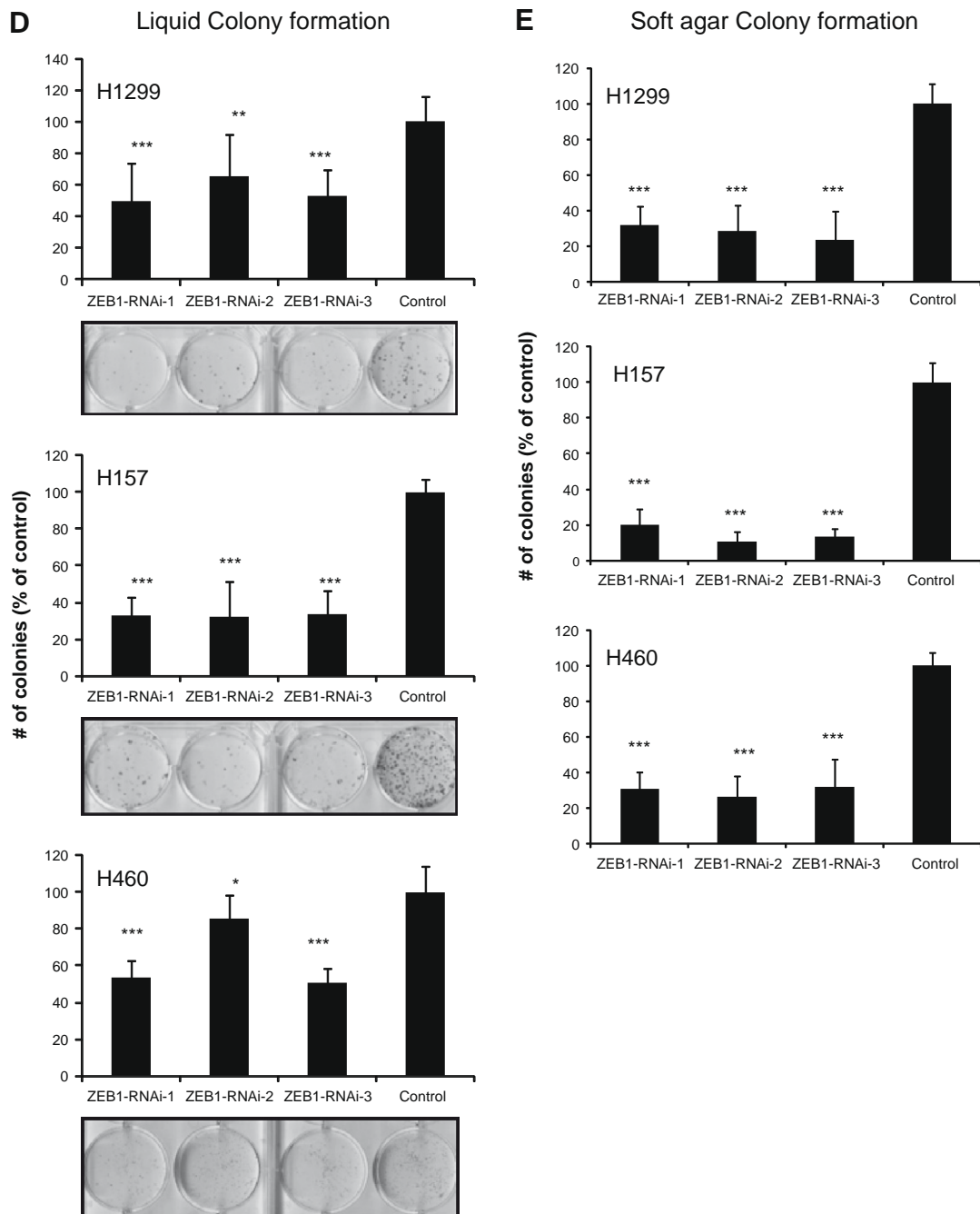


Fig. 2 (continued)

premature senescence, suggesting that *ZEB1* may function as inhibitor of senescence [29]. Nevertheless, we did not see any increase in the number of cells exhibiting morphologic changes or β -galactosidase staining suggestive of senescence, in any of cells transfected with *ZEB1* knockdown oligos (data not shown).

4. Discussion

In the present study, we have shown that *ZEB1* expression was correlated with the mesenchymal phenotype in

NSCLC, and that its depletion with RNA interference suppressed anchorage-independent growth, thus providing a rationale for developing therapeutics targeting *ZEB1* function in lung cancer. Consistent with reports in other cancers, we also found that miR-200c and miR-205 expression were inversely correlated with *ZEB1* expression in NSCLCs, suggesting potential use as therapy targeting *ZEB1*.

Consistent with the literature [30], we found that *EGFR* mutations were associated with epithelial characteristics

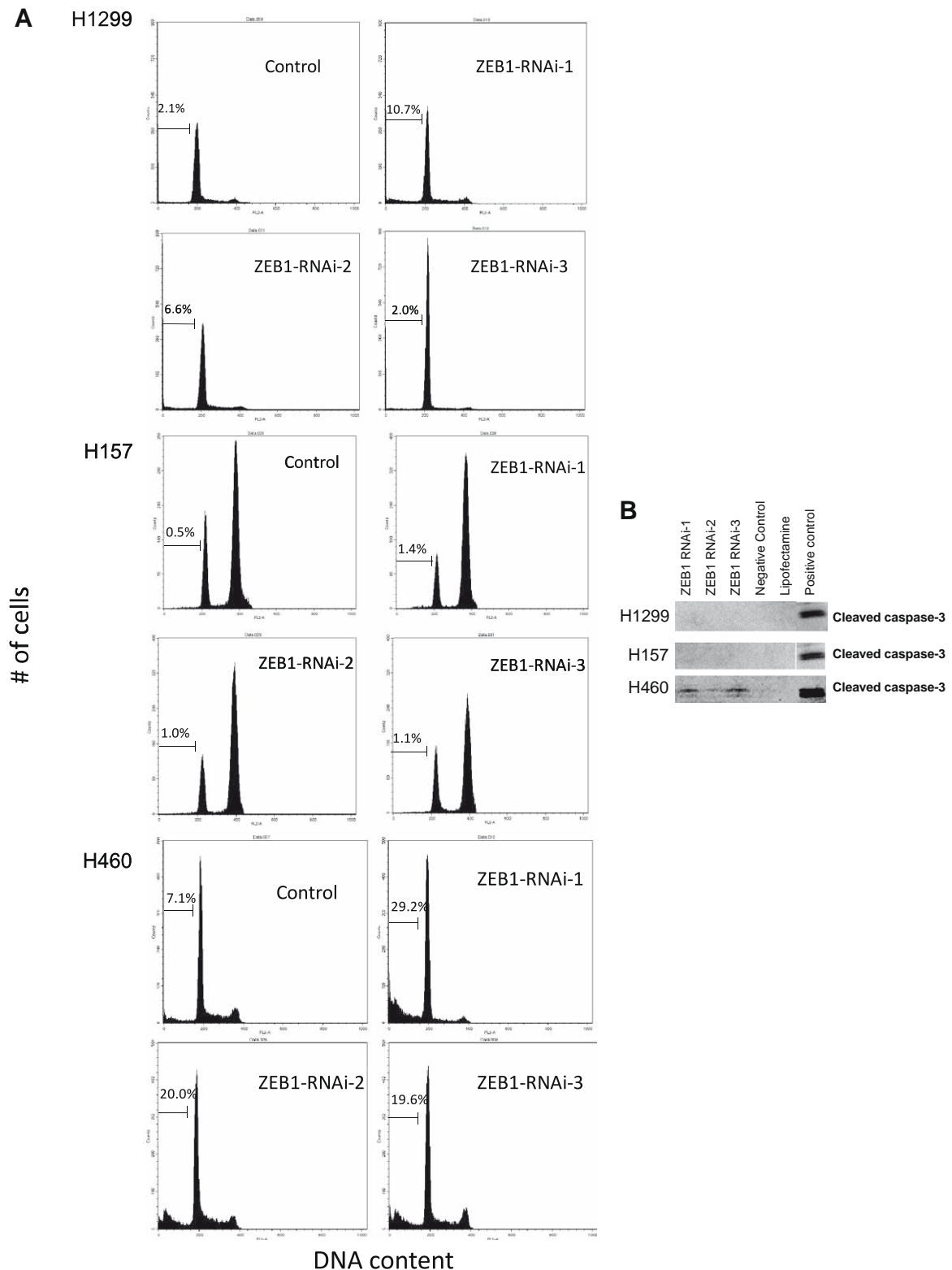


Fig. 3. *ZEB1* knockdown induces apoptosis in H460. (A) Flow cytometry analysis for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. Forty-eight hours after the transfection cells were stained with propidium iodide and 20,000 cells were analyzed for cell cycle profiling. Representative results in three independent experiments are shown. (B) Western blot of cleaved caspase-3 for lung cancer cell lines transfected with *ZEB1* knockdown or control oligos. ACC-MESO-1 mesothelioma cell line treated with cisplatin was used as a positive control.

(E-cadherin expression) in NSCLCs. The literature study showed that *EGFR* mutant genotype is an independent predictor for the epithelial phenotype in logistic regression analysis, excluding the possibility that observed high frequency of the epithelial phenotype in *EGFR* mutants was due to other factors that may be associated with or may cause epithelial phenotype. Such factors include early stage disease and non- or low smoking history. The finding that *EGFR* mutant cells frequently have an epithelial phenotype seems to be unexpected because signaling pathways activated by *EGFR* including RAS, AKT, and SRC pathways can induce EMT [31]. It is possible that *EGFR* mutant cells are resistant to EMT-inducing signals. Alternatively, it is also possible that mutant cells specifically express genes that retain cells to the epithelial state. It will be important to further elucidate the molecular mechanisms leading to epithelial phenotype in *EGFR* mutant cells because molecules involved in such mechanisms could serve as therapeutic targets.

We found that in NSCLC cell lines *ZEB1* expression (in contrast to other master EMT genes) was significantly correlated with both *Vimentin* and *E-cadherin* expression. In addition, we saw that *ZEB1* expression was very strongly correlated with *Vimentin* expression in primary tumors. Two previous studies also reported that *ZEB1* expression was most significantly correlated with E-cadherin expression in lung cancer among several EMT-inducing genes [18,32]. Although *ZEB1* has been shown to directly repress transcription of *E-cadherin* in several different types of cells [33,34], it remains unknown whether *ZEB1* up-regulates *Vimentin* expression. In our *ZEB1* knockdown experiments, we did not see a significant down-regulation of *Vimentin* in any of the three lines studied, suggesting that *ZEB1* does not directly up-regulate *Vimentin* expression. However, the observed strong correlation between *ZEB1* and *Vimentin* expression in both NSCLC cell lines and tumors suggests that *ZEB1* may indirectly up-regulate *Vimentin* possibly through some of its downstream targets. Collectively, our results, along with those of others, suggest a dominant role of *ZEB1* in maintaining the mesenchymal phenotype of lung cancer.

ZEB1 knockdown induced *E-cadherin* mRNA expression in H157 and H460 but not in H1299. There is one possible explanation for this. In H1299, the promoter region of *E-cadherin* is known to be heavily methylated, causing its silenced expression [35]. Thus, it is possible that transient knockdown of *ZEB1* was unable to overcome methylation-mediated gene silencing of *E-cadherin* in H1299, resulting in the unchanged *E-cadherin* mRNA. Long-term *ZEB1* knockdown could reverse methylation of *E-cadherin*, and thus it would be interesting to see the effects of stable *ZEB1* knockdown on methylation status of *E-cadherin* in H1299.

ZEB1 knockdown inhibited NSCLC growth most dramatically noted in soft agar colony formation assay. Since the ability of transformed cells to grow under anchorage-independent condition is the most reliable predictor for tumorigenicity and metastatic potential, this strong growth inhibitory effect of *ZEB1* knockdown in anchorage-independent condition suggests that *ZEB1* expression contributes to maintaining aggressive phenotype of lung cancer cells.

Studies have shown that *Twist*, *Snail* and *SIP1* show anti-apoptotic effects when ectopically expressed [36–38]. Nevertheless, to our knowledge, only one study has demonstrated anti-apoptotic effect of a master EMT gene in a gene-silencing experiment; the study demonstrated that *Twist* knockdown induced apoptosis in human neuroblastoma cells [39]. We also observed that *ZEB1* knockdown induced apoptosis in H460. These results including ours demonstrate that in some cancer cells survival signals are dependent on expression of master EMT genes such as *Twist* and *ZEB1*. Cancer cells are thought to acquire the ability to evade apoptosis in early stage of progression, and thus these findings suggest that master EMT genes may play an important role not only in late stage but also in early stage of carcinogenesis.

In conclusion, we have shown that *ZEB1* expression is very well correlated with the mesenchymal phenotype of NSCLC, and that its removal induces significant growth inhibition in NSCLC partially through induction of apoptosis. These results suggest that *ZEB1* is a promising therapeutic target for lung cancer.

5. Conflict of interest

We have no conflict of interest to be disclosed.

Acknowledgements

This work was supported by Grant-in-Aid for Scientific Research (C) 20590919 (to M.S.), Grant-in-Aid for Scientific Research (C) 20590918 (to M.K.), and Grant-in-Aid for Scientific Research (B) 21390257 (to Y.H.) from Japan Society for the Promotion of Science, Global COE program in Nagoya University Graduate School of Medicine funded by Japan's Ministry of Education, Culture, Sports, Science and Technology, NCI Special Program of Research Excellence in Lung Cancer (SPORE P50CA70907), and DOD PROSPECT (to JDM and AFG). We thank Dr. David S. Shames for his helpful comments and discussion on this study.

References

- [1] J.J. Christiansen, A.K. Rajasekaran, Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis, *Cancer Res.* 66 (2006) 8319–8326.
- [2] M.W. Klymkowsky, P. Savagner, Epithelial–mesenchymal transition: a cancer researcher's conceptual friend and foe, *Am. J. Pathol.* 174 (2009) 1588–1593.
- [3] C. Vandewalle, F. Van Roy, G. Berx, The role of the ZEB family of transcription factors in development and disease, *Cell. Mol. Life Sci.* 66 (2009) 773–787.
- [4] E. Batlle, E. Sancho, C. Franci, D. Dominguez, M. Monfar, J. Baulida, A. Garcia De Herreros, The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells, *Nat. Cell Biol.* 2 (2000) 84–89.
- [5] A. Cano, M.A. Perez-Moreno, I. Rodrigo, A. Locascio, M.J. Blanco, M.G. del Barrio, F. Portillo, M.A. Nieto, The transcription factor snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression, *Nat. Cell Biol.* 2 (2000) 76–83.
- [6] J. Yang, S.A. Mani, J.L. Donaher, S. Ramaswamy, R.A. Itzykson, C. Come, P. Savagner, I. Gitelman, A. Richardson, R.A. Weinberg, Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis, *Cell* 117 (2004) 927–939.
- [7] A. Miyoshi, Y. Kitajima, S. Kido, T. Shimonishi, S. Matsuyama, K. Kitahara, K. Miyazaki, Snail accelerates cancer invasion by upregulating MMP expression and is associated with poor

- prognosis of hepatocellular carcinoma, *Br. J. Cancer* 92 (2005) 252–258.
- [8] Z. Yan-Qi, G. Xue-Yan, H. Shuang, C. Yu, G. Fu-Lin, B. Fei-Hu, S. Shi-Ren, W. Xu-Feng, D. Jie, F. Dai-Ming, Expression and significance of TWIST basic helix-loop-helix protein over-expression in gastric cancer, *Pathology* 39 (2007) 470–475.
 - [9] F. Vesuna, P. van Diest, J.H. Chen, V. Raman, Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer, *Biochem. Biophys. Res. Commun.* 367 (2008) 235–241.
 - [10] F. Valdes-Mora, T. Gomez del Pulgar, E. Bandres, P. Cejas, A. Ramirez de Molina, R. Perez-Palacios, D. Gallego-Ortega, M.A. Garcia-Cabezas, E. Casado, J. Larrauri, M. Nistal, M. Gonzalez-Baron, J. Garcia-Foncillas, J.C. Lacal, TWIST1 overexpression is associated with nodal invasion and male sex in primary colorectal cancer, *Ann. Surg. Oncol.* 16 (2009) 78–87.
 - [11] D.M. Parkin, F. Bray, J. Ferlay, P. Pisani, Global cancer statistics, 2002, *CA Cancer J. Clin.* 55 (2005) 74–108.
 - [12] M. Sato, D.S. Shames, A.F. Gazdar, J.D. Minna, A translational view of the molecular pathogenesis of lung cancer, *J. Thorac. Oncol.* 2 (2007) 327–343.
 - [13] Y. Sekido, K.M. Fong, J.D. Minna, Molecular genetics of lung cancer, *Annu. Rev. Med.* 54 (2003) 73–87.
 - [14] L. Prudkin, D.D. Liu, N.C. Ozburn, M. Sun, C. Behrens, X. Tang, K.C. Brown, B.N. Bekele, C. Moran, Wistuba II, Epithelial-to-mesenchymal transition in the development and progression of adenocarcinoma and squamous cell carcinoma of the lung, *Mod. Pathol.* 22 (2009) 668–678.
 - [15] S. Spaderna, O. Schmalhofer, M. Wahlbuhl, A. Dimmler, K. Bauer, A. Sultan, F. Hlubek, A. Jung, D. Strand, A. Eger, T. Kirchner, J. Behrens, T. Brabletz, The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer, *Cancer Res.* 68 (2008) 537–544.
 - [16] T. Arumugam, V. Ramachandran, K.F. Fournier, H. Wang, L. Marquis, J.L. Abbruzzese, G.E. Gallick, C.D. Logsdon, D.J. McConkey, W. Choi, Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer, *Cancer Res.* 69 (2009) 5820–5828.
 - [17] Y. Li, T.G. VandenBoom II, D. Kong, Z. Wang, S. Ali, P.A. Philip, F.H. Sarkar, Up-regulation of miR-200 and let-7 by natural agents leads to the reversal of epithelial-to-mesenchymal transition in gemcitabine-resistant pancreatic cancer cells, *Cancer Res.* 69 (2009) 6704–6712.
 - [18] T. Ohira, R.M. Gemmill, K. Ferguson, S. Kusy, J. Roche, E. Brambilla, C. Zeng, A. Baron, L. Bemis, P. Erickson, E. Wilder, A. Rustgi, J. Kitajewski, E. Gabrielson, R. Bremnes, W. Franklin, H.A. Drabkin, WNT7a induces E-cadherin in lung cancer cells, *Proc. Natl. Acad. Sci. USA* 100 (2003) 10429–10434.
 - [19] M. Dohadwala, S.C. Yang, J. Luo, S. Sharma, R.K. Batra, M. Huang, Y. Lin, L. Goodlick, K. Krysan, M.C. Fishbein, L. Hong, C. Lai, R.B. Cameron, R.M. Gemmill, H.A. Drabkin, S.M. Dubinett, Cyclooxygenase-2-dependent regulation of E-cadherin: prostaglandin E(2) induces transcriptional repressors ZEB1 and snail in non-small cell lung cancer, *Cancer Res.* 66 (2006) 5338–5345.
 - [20] J. Clahaut, R.M. Gemmill, V.A. Potiron, S. Ait-Si-Ali, J. Imbert, H.A. Drabkin, J. Roche, ZEB-1, a repressor of the semaphorin 3F tumor suppressor gene in lung cancer cells, *Neoplasia* 11 (2009) 157–166.
 - [21] J. Gandhi, J. Zhang, Y. Xie, J. Soh, H. Shigematsu, W. Zhang, H. Yamamoto, M. Peyton, L. Girard, W.W. Lockwood, W.L. Lam, M. Varella-Garcia, J.D. Minna, A.F. Gazdar, Alterations in genes of the EGFR signaling pathway and their relationship to EGFR tyrosine kinase inhibitor sensitivity in lung cancer cell lines, *PLoS One* 4 (2009) e4576.
 - [22] N. Usami, T. Fukui, M. Kondo, T. Taniguchi, T. Yokoyama, S. Mori, K. Yokoi, Y. Horio, K. Shimokata, Y. Sekido, T. Hida, Establishment and characterization of four malignant pleural mesothelioma cell lines from Japanese patients, *Cancer Sci.* 97 (2006) 387–394.
 - [23] T. Yokoyama, M. Kondo, Y. Goto, T. Fukui, H. Yoshioka, K. Yokoi, H. Osada, K. Imaizumi, Y. Hasegawa, K. Shimokata, Y. Sekido, EGFR point mutation in non-small cell lung cancer is occasionally accompanied by a second mutation or amplification, *Cancer Sci.* 97 (2006) 753–759.
 - [24] M. Sato, M.B. Vaughan, L. Girard, M. Peyton, W. Lee, D.S. Shames, R.D. Ramirez, N. Sunaga, A.F. Gazdar, J.W. Shay, J.D. Minna, Multiple oncogenic changes (K-RAS(V12), p53 knockdown, mutant EGFRs, p16 bypass, telomerase) are not sufficient to confer a full malignant phenotype on human bronchial epithelial cells, *Cancer Res.* 66 (2006) 2116–2128.
 - [25] G.J. Hurteau, J.A. Carlson, S.D. Spivack, G.J. Brock, Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin, *Cancer Res.* 67 (2007) 7972–7976.
 - [26] U. Burk, J. Schubert, U. Wellner, O. Schmalhofer, E. Vincan, S. Spaderna, T. Brabletz, A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells, *EMBO Rep.* 9 (2008) 582–589.
 - [27] P.A. Gregory, A.G. Bert, E.L. Paterson, S.C. Barry, A. Tsykin, G. Farshid, M.A. Vadas, Y. Khew-Goodall, G.J. Goodall, The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1, *Nat. Cell Biol.* 10 (2008) 593–601.
 - [28] S.M. Park, A.B. Gaur, E. Lengyel, M.E. Peter, The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2, *Genes Dev.* 22 (2008) 894–907.
 - [29] Y. Liu, S. El-Naggar, D.S. Darling, Y. Higashi, D.C. Dean, Zeb1 links epithelial-mesenchymal transition and cellular senescence, *Development* 135 (2008) 579–588.
 - [30] Q.F. Deng, C.C. Zhou, C.X. Su, Clinicopathological features and epidermal growth factor receptor mutations associated with epithelial-mesenchymal transition in non-small cell lung cancer, *Respirology* 14 (2009) 371–376.
 - [31] L. Larue, A. Bellacosa, Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways, *Oncogene* 24 (2005) 7443–7454.
 - [32] S.E. Witta, R.M. Gemmill, F.R. Hirsch, C.D. Coldren, K. Hedman, L. Ravel, B. Helfrich, R. Dziadziuszko, D.C. Chan, M. Sugita, Z. Chan, A. Baron, W. Franklin, H.A. Drabkin, L. Girard, A.F. Gazdar, J.D. Minna, P.A. Bunn Jr., Restoring E-cadherin expression increases sensitivity to epidermal growth factor receptor inhibitors in lung cancer cell lines, *Cancer Res.* 66 (2006) 944–950.
 - [33] M.L. Grootenclaes, S.M. Frisch, Evidence for a function of CtBP in epithelial gene regulation and anoikis, *Oncogene* 19 (2000) 3823–3828.
 - [34] S. Guaita, I. Puig, C. Franci, M. Garrido, D. Dominguez, E. Battle, E. Sancho, S. Dedhar, A.G. De Herreros, J. Baulida, Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression, *J. Biol. Chem.* 277 (2002) 39209–39216.
 - [35] M. Suzuki, N. Sunaga, D.S. Shames, S. Toyooka, A.F. Gazdar, J.D. Minna, RNA interference-mediated knockdown of DNA methyltransferase 1 leads to promoter demethylation and gene re-expression in human lung and breast cancer cells, *Cancer Res.* 64 (2004) 3137–3143.
 - [36] M. Kajita, K.N. McClinic, P.A. Wade, Aberrant expression of the transcription factors snail and slug alters the response to genotoxic stress, *Mol. Cell Biol.* 24 (2004) 7559–7566.
 - [37] S. Vega, A.V. Morales, O.H. Ocana, F. Valdes, I. Fabregat, M.A. Nieto, Snail blocks the cell cycle and confers resistance to cell death, *Genes Dev.* 18 (2004) 1131–1143.
 - [38] A.E. Sayan, T.R. Griffiths, R. Pal, G.J. Browne, A. Ruddick, T. Yagci, R. Edwards, N.J. Mayer, H. Qazi, S. Goyal, S. Fernandez, K. Straatman, G.D. Jones, K.J. Bowman, A. Colquhoun, J.K. Mellon, M. Kriaevska, E. Tulchinsky, SIP1 protein protects cells from DNA damage-induced apoptosis and has independent prognostic value in bladder cancer, *Proc. Natl. Acad. Sci. USA* (2009).
 - [39] S. Valsesia-Wittmann, M. Magdeleine, S. Dupasquier, E. Garin, A.C. Jallas, V. Combaret, A. Krause, P. Leissner, A. Puisieux, Oncogenic cooperation between H-Twist and N-Myc overrides failsafe programs in cancer cells, *Cancer Cell* 6 (2004) 625–630.



Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan



Histologic features of low- and intermediate-grade neuroendocrine carcinoma (typical and atypical carcinoid tumors) of the lung

Koji Tsuta^{a,d}, Maria G. Raso^b, Neda Kalhor^a, Diane D. Liu^c, Ignacio I. Wistuba^{a,b}, Cesar A. Moran^{a,*}

^a Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

^b Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

^c Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

^d Division of Clinical Laboratory, National Cancer Center Hospital, Tokyo, Japan

ARTICLE INFO

Article history:

Received 10 November 2009

Received in revised form 5 February 2010

Accepted 1 April 2010

Keywords:

Typical carcinoid

Atypical carcinoid

Mitosis

Necrosis

Recurrence-free survival

Neuroendocrine carcinoma

ABSTRACT

Background: Determining the differential diagnosis between typical (TCs) and atypical carcinoid tumors (ACs) is imperative, as the distinction between TCs and ACs is currently based on histologic criteria that are not always correlated with the unfavorable clinical outcomes.

Patients and methods: We conducted a retrospective study of patients who were diagnosed with carcinoid tumors between 1990 and 2005 at M. D. Anderson Cancer Center. We reviewed the slides for the following pathologic features: infiltrative growth; pleural, blood, or lymphatic vessel invasion; tumor stroma; presence of active fibroblastic proliferation; chromatin pattern; presence of nucleolus; and nuclear pleomorphism. We also evaluated the necrotic patterns. Finally, we evaluated three methods for calculating the number of mitoses: randomly selected, the most mitotically active in 10 high-power fields (HPFs), or overall mean mitotic count.

Results: Our cohort consisted of 80 patients (68 with TCs and 12 with ACs). Older age ($P=0.002$), pathologic stage III or IV disease ($P=0.04$), active fibroblastic proliferation ($P=0.041$), and comedo-like necrosis ($P=0.001$) were significantly associated with tumor recurrence or patient's death. Among the three mitotic counting methods, the overall mean number of mitoses was significantly correlated with recurrence-free survival ($P<0.0001$). Our criteria for distinguishing AC from TC included the presence of comedo-like necrosis and/or an overall mean number of mitoses $\geq 0.2/\text{HPF}$.

Conclusions: Using an overall mean number in counting mitoses and detecting comedo-like necrosis is important for classifying lung carcinoid tumors.

© 2010 Published by Elsevier Ireland Ltd.

1. Introduction

Bronchopulmonary neuroendocrine tumors comprise about 20% of all lung carcinomas and represent a spectrum of tumors arising from neuroendocrine cells. The different types of bronchopulmonary neuroendocrine carcinomas share structural, morphologic, immunohistochemical, and ultrastructural features, and they are separated into four subgroups in increasing order of biologic aggressiveness: low-grade (typical carcinoids; TCs), intermediate-grade (atypical carcinoids; ACs), and high-grade (large-cell neuroendocrine carcinomas; LCNECs), or (small-cell lung carcinomas). TCs and ACs account for approximately 1–2% of all primary lung carcinomas [1,2].

Although TCs are generally regarded as low-grade carcinomas, approximately 10–23% [3] of cases metastasize to the regional lymph nodes at presentation, with the 5-year overall survival rates ranging from 82% to 100% for patients with TCs [4,5]. In contrast, approximately 40–50% of ACs metastasize to the regional lymph nodes at presentation, with the 5-year overall survival rates ranging from 25% to 78% for patients with ACs [2,5–9].

Conventional carcinoid tumors are composed of homogeneous cellular proliferation with an organoid and trabecular structure. Tumor cells are composed of small- to medium-sized, round to polygonal cells with a scant to moderate amount of eosinophilic cytoplasm and centrally located round to oval nuclei with fine granular chromatin. Histologic atypia, such as nuclear atypia, prominent nucleoli, and pleomorphism, is more common in AC than TC, but these features are sometimes observed in TC [6,7,10]. Therefore, the unfavorable clinical outcomes and atypical histologic features associated with these lung carcinoid tumors are unreliable for distinguishing TCs from ACs.

* Corresponding author at: Department of Pathology, Unit 85, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. Tel.: +1 713 792 8134; fax: +1 713 745 3740.

E-mail address: CesarMoran@mdanderson.org (C.A. Moran).

Arrigoni et al. [11] first defined ACs as carcinoid tumors with (1) 1 mitosis/1–2 high-power fields (HPFs) or 5–10 mitoses/10 HPF; (2) necrosis; (3) pleomorphism, hyperchromasia, or an abnormal nuclear/cytoplasmic ratio; and (4) areas of increased cellularity with disorganization. In 1998, Travis et al. [6] proposed that LCNECs be classified in a separate category from ACs, as this poorly differentiated high-grade carcinoma was characterized by a neuroendocrine appearance under light microscopy. These criteria have been applied to the latest World Health Organization classification [10]. With the establishment of diagnostic criteria for LCNEC, the definition of AC was then restricted to tumors with a mitotic rate of 2–10 mitoses per 2 mm² (10 HPF) or the presence of necrosis.

Differentiating ACs from TCs or LCNECs is clinically important because the treatment modalities and prognoses for these types of tumors are different. However, the distinction between ACs and TCs is currently based only on the histologic evaluation of mitotic count and necrosis. In this study, we sought to retrospectively evaluate the various histologic features of ACs and TCs and relate these to their outcomes (i.e., tumor recurrence, patient death, or lymph node metastasis). Furthermore, we revised the relevance of the patterns of necrosis and mitotic counting methods.

2. Materials and methods

2.1. Case selection

We retrospectively reviewed the tumor specimens from patients who were surgically resected and diagnosed with TC or AC between 1990 and 2005. We obtained the tumor specimens from cases deposited in the files of the University of Texas M. D. Anderson Cancer Center (Houston, TX, USA) Lung SPORE Tissue Bank, which was approved by the institutional review board. We reviewed the pathologic records of the specimens and all available hematoxylin and eosin (HE)-stained slides, some special stains, and the immunohistochemical and/or ultrastructural analyses available. We collected the patients' clinical information, including age, gender, smoking history, presenting symptoms, treatment modalities, site of any tumor recurrence, duration of recurrence or survival, tumor location (which lobe and whether central or not), maximum tumor size (in cm), and pathologic disease stage (p-stage).

2.2. Histologic examination

All available HE-stained slides for each case were examined by two pathologists (K.T. and C.A.M.); each pathologist was blinded to the clinical details of each patient. We evaluated the pathologic features of tumor invasiveness, including the infiltrative growth of the adjacent normal architecture, presence of pleural invasion, and presence of blood or lymphatic vessel invasion. The stroma between the tumor nests was mainly composed of capillary or fibrous tissue. We also observed active fibroblastic proliferation [12]. We then evaluated the histologic features of the tumor cells, including the chromatin pattern (i.e., fine granular or coarse) and the presence or absence of nucleolus and nuclear pleomorphism (i.e., whether there was more than a 3-fold variation in nuclear size that did not include the presence of giant cells).

We then evaluated the necrotic pattern, such as the presence or absence of comedo-like necrosis (necrosis present within the center of tumor nest) and coagulative necrosis (necrosis involving one or more whole tumor nests and/or forming bridging necrosis) (Fig. 1A). We also evaluated the histologic features suggesting tumor damage, such as pyknotic cells (i.e., foci of dyscohesive cells) with somewhat pyknotic nuclei and condensed, shrunken-appearing cytoplasm without karyorrhexis, eosinophilic debris, or apoptosis (Fig. 1B), cholesterol cleft, stromal hemorrhage, and psammoma bodies.

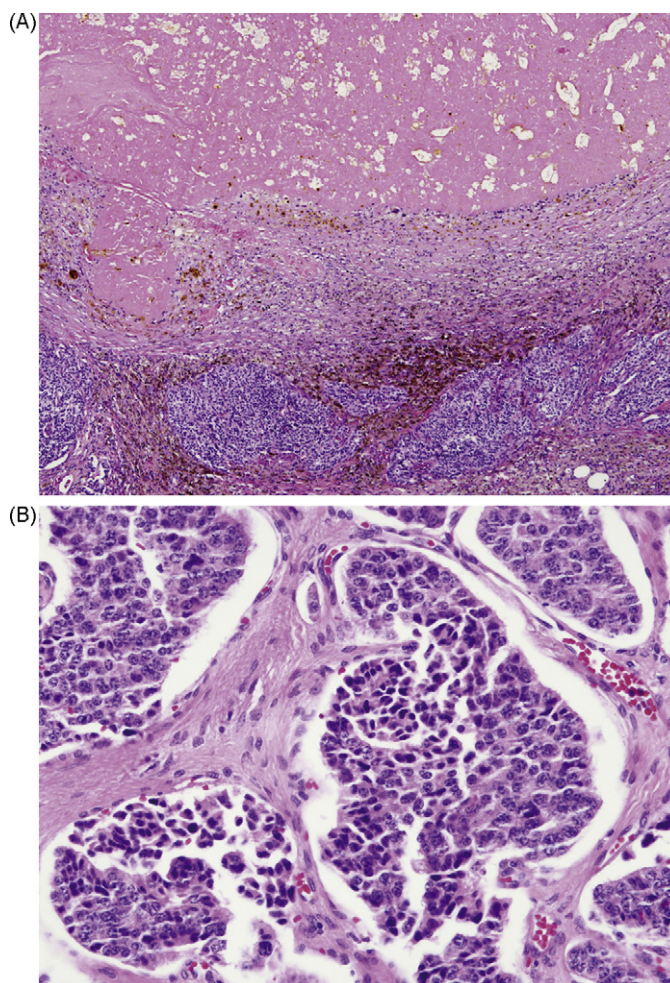


Fig. 1. (A) Massive coagulative necrosis surrounded by thick fibrous tissue with a hemosiderin deposition. The outermost portion of the necrosis consisted of tumor cells. (B) Foci of dyscohesive cells with somewhat pyknotic nuclei and condensed, shrunken-appearing cytoplasm without karyorrhexis, eosinophilic debris, or apoptosis observed in tumor nests.

or apoptosis (Fig. 1B), cholesterol cleft, stromal hemorrhage, and psammoma bodies.

2.3. The mitotic counting method

After all slides were reviewed on an Olympus CX31 microscope (Olympus; Tokyo, Japan), the mitoses were counted on one representative slide. This microscope's standard field of view number is 20 (0.2 mm²); therefore, an HPF magnification of 400× equals to 0.2 mm². To determine the most suitable mitotic counting method, we studied three methods: (1) the number of mitoses in a randomly selected 10-HPF area counted; (2) the number of mitoses in the most mitotically active 10-HPF areas (so called hot spots) counted; or (3) the mean number of mitoses in the whole section calculated (the number of mitoses divided by the number of HPFs in the whole section). To avoid overcounting the number of mitoses at the telophase, two adjoining mitotic figures were counted as one mitosis.

2.4. Statistical analysis

We performed all statistical analyses using SPSS version 12.0 software for Windows (SPSS; Chicago, IL). We used the Wilcoxon rank-sum test to analyze the continuous variables and Fisher's

exact test to analyze lymph node metastasis. We calculated the recurrence-free survival (RFS) curves using the Kaplan–Meier method, and we compared the curves using the log-rank test. We considered a *P*-value of ≤ 0.05 to be significant in two tailed analysis.

3. Results

3.1. Clinical features

We reviewed the tumor specimens from 88 patients with an original diagnosis of lung carcinoid tumors. We excluded 5 patients who received therapy preoperatively (4 who received neoadjuvant chemotherapy and 1 who received YAG laser ablation). Furthermore, we excluded 3 patients who had high mitotic activity and massive necrosis; their diagnosis was then changed to LCNEC based on the latest World Health Organization criteria [10]. Therefore, our final cohort consisted of 80 patients with lung carcinoid tumors, including 68 patients with TCs and 12 with ACs. The patients' mean age (at the time of diagnosis) was 54.3 years (range, 19–80 years). The presenting symptoms were documented in 27 out of 77 patients. Smoking history was documented in 76 cases; 32 patients were never smokers, 32 were former smokers, and 12 were present smokers. The tumor was located in the right-upper lobe in 11 patients, the middle lobe in 10, the right-lower lobe in 26, the left-upper lobe in 13, and the left-lower lobe in 19, and 1 patient had a hilar lesion. The tumor was located in the central portion of the lung in 49 patients or in the intermediate or peripheral portion in 31 patients. A preoperative biopsy and/or cytologic diagnostic procedure were carried out in 69 patients, including the diagnoses were not tumorous lesions.

The tumor sizes ranged from 1.0 to 13.0 cm (mean, 2.96 cm). Lymph node status was recorded in 79 patients, and metastasis was observed in 17. Sixty-one cases had p-stage I disease, 6 had p-stage II disease, 11 had p-stage III disease, 1 had p-stage IV disease, and 1 case was unstaged.

Tumor recurrence was observed in 6 of 80 patients (2 cases of liver recurrence and 1 case each of kidney, brain, chest wall, and subclavicular lymph node recurrence). The mean follow-up time for all 80 patients was 68 months (range, 1–197 months), with 74 still alive at the time of this report. Only 1 patient died of a tumor progression, and 5 died of other causes.

3.2. Histologic factors

We reviewed an average of 4.7 (range, 1–16) HE-stained slides for each case. Infiltrative growth was observed in 47 cases. Blood-vessel, lymph-vessel, and pleural invasion were observed in 34 cases, 29 cases, and 1 case, respectively. Capillary rich stroma was observed in 57 cases, and fibrous rich stroma was observed in 23 cases. Active fibroblastic proliferation was observed in 38 cases (Fig. 2A). We found a fine chromatin pattern in 47 cases and a coarse chromatin pattern in 33 cases. We also observed a conspicuous nucleolus in 15 cases and nuclear pleomorphism in 13 cases. We found other unusual histologic features of the tumors, such as bone formation in 16 cases and granulomatous inflammation in 2 cases.

We found comedo-like necrosis in 5 cases and coagulative necrosis in 10 cases (Fig. 2B), and we found both comedo-like and coagulative necrosis in 2 cases. All but 1 patient with coagulative necrosis had undergone preoperative biopsy. Pyknotic cells, cholesterol cleft, hemorrhage, and psammoma bodies were observed in 54 cases, 4 cases, 29 cases, and 13 cases, respectively.

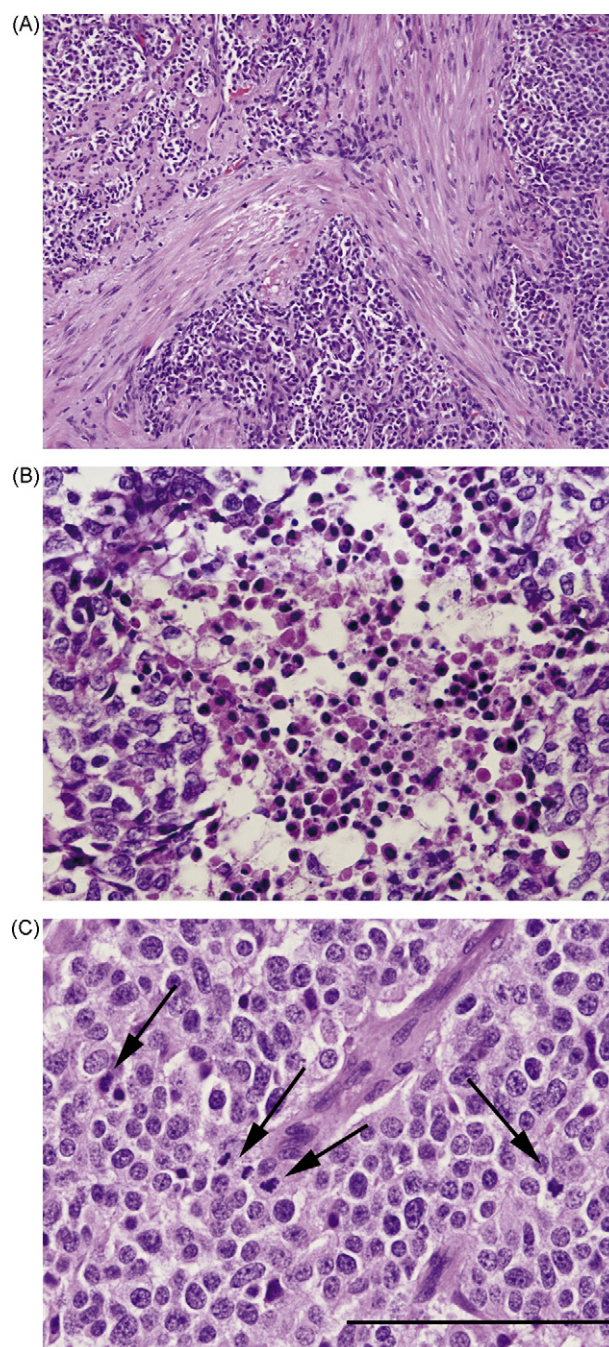


Fig. 2. (A) Bundles of spindle-shaped fibroblasts are shown between tumor nests (magnification, 10 \times). (B) The tumor had a large solid nest with comedo-like necrosis (magnification, 20 \times). (C) Although this case showed 5 mitoses in 1 high-power field (HPF), this photomicrograph area is smaller than 1 HPF (0.2 mm²) and corresponded to a 0.134 mm² area. Because this image was captured using the Olympus DP-70 system (Tokyo, Japan), this photomicrograph showed 4 mitotic figures, indicated by the arrows (2 in prophase, 1 in metaphase, and 1 in telophase). This case's mitotic count in the most mitotically active area was 15 mitoses/10 HPF, but the average mitosis was 0.18 mitoses/HPF. This patient was alive at 45 months with no evidence of tumor recurrence (bar = 200 μ m).

3.3. Mitotic count

We reviewed a mean of 341.6 HPF (range, 44–762 HPF) per case. The number of mitoses ranged from 0 to 115, with a mean of 10.9 per case. The number of mitoses in 1 HPF exceeded 2 in 26 cases, and the highest recorded mitotic count in 1 HPF was 5 (Fig. 2C). The mean number of mitoses in the randomly selected 10-HPF

areas was 0.45 mitoses/10 HPF (range, 0–6 mitoses/10 HPF), and the mean number of mitoses in the most mitotically active 10-HPF areas was 5.73 mitoses/10 HPF (range, 0–25 mitoses/10 HPF). The overall mean number of mitoses was 0.04 mitoses/1 HPF (range, 0–0.32 mitoses/1 HPF).

Both the presence of comedo-like necrosis and an overall mean number of mitoses of ≥ 0.2 mitoses/1 HPF were observed in 3 cases. These cases corresponded to 60% of cases bearing comedo-like necrosis and 60% of cases having an overall mean number of mitoses ≥ 0.2 mitoses/1 HPF.

3.4. Clinicopathologic features based on unfavorable outcomes

We next evaluated whether there was an association between the patients' clinicopathologic factors and tumor recurrence, patient death, and/or lymph node metastasis (Table 1). An older age ($P = 0.002$), p-stage III or IV disease ($P = 0.0029$), pleural invasion ($P < 0.0001$), and the presence of active fibroblasts ($P = 0.0184$) were significantly associated with tumor recurrence or patient death. All patients with tumor specimens that exhibited an active fibroblastic proliferation had undergone a preoperative biopsy and/or cytologic diagnostic procedure. Blood-vessel invasion ($P = 0.0417$), lymph-vessel invasion ($P = 0.0017$), coarse chromatin pattern ($P = 0.0304$), conspicuous nucleolus ($P = 0.0085$), and nuclear pleomorphism ($P = 0.0019$) were significantly associated with lymph node metastasis.

3.5. Evaluation of necrotic patterns based on the patients' unfavorable outcomes and RFS

We then sought to determine if there was a correlation between the tumor specimens' necrotic patterns and histologic features suggesting tumor damage and the patients' unfavorable outcomes (Table 2). Only the presence of comedo-like necrosis was significantly associated with tumor recurrence or patient death ($P < 0.0001$).

3.6. Evaluation of mitotic count based on the patients' unfavorable outcomes, RFS, and OS

We investigated whether there was a correlation between the results of three mitotic counting methods and the patients' unfavorable outcomes and RFS (Table 3). For specimens in which the number of mitoses in a randomly selected 10-HPF area was counted, the presence of 2–10 mitoses/10 HPF appeared to be significantly associated with tumor recurrence or patient death ($P = 0.003$). In addition, the 5-year RFS rates for the specimens with 0–1 mitoses/10 HPF and 2–10 mitoses/10 HPF were 94.2% and 58.3%, respectively (Fig. 3A). However, when we individualized each case and correlated the number of mitotic figures with the 5-year RFS rates, we obtained the following information: 0 mitoses = 95.1%; 1 mitotic figure = 87.5%; 2 mitotic figures = 75.0%; 3 mitotic figures = 100%. Based on these data, we analyzed changing the cut off number of mitosis from 0–3 to 4–10. The 5-year RFS rates for the specimens with 0–3 mitoses/10 HPF and 4–10 mitoses/10 HPF were 93.1% and 0%, respectively ($P = 0.002$).

For specimens in which the number of mitoses in the most mitotically active 10-HPF area was counted, this method showed tendency with tumor recurrence or patient death ($P = 0.0575$). The 5-year RFS rates for specimens with 0–1 mitoses/10 HPF, 2–10 mitoses/10 HPF, and more than 11 mitoses/10 HPF were 95.5%, 94.6%, and 75.2%, respectively ($P = 0.0575$; Fig. 3B). Once again, when we individualized each case and correlated the number of mitotic figures with the 5-year RFS rates, we obtained the following information: 0 mitoses = 100%; 1 mitotic figure = 90.0%; 2 mitotic figures = 100%; and 3 mitotic figures = 100%. Based on these

data, we analyzed changing the cut off number of mitosis from 0–3 to 4 or more. The 5-year RFS rates for the specimens with 0–3 mitoses/10 HPF and 4 or more mitoses/10 HPF were 97.4% and 83.4%, respectively ($P = 0.024$).

For specimens in which the overall mean number of mitoses was calculated, the value of ≥ 0.2 mitoses per 1 HPF was significantly associated with tumor recurrence or patient death ($P < 0.0001$). The 5-year RFS rates for specimens with an overall mean mitosis < 0.2 and a mean mitosis ≥ 0.2 were 95.9% and 0%, respectively (Fig. 3C). Multivariate analysis was not performed because of the small number of events (recurrence and death) in our patient population. Based on the 5-year RFS rates, we considered that the overall mean mitotic counting method was the most useful. Furthermore, we also found that the criteria on the number of mitosis should be selected based on the counting methods (0–3 and 4–10 in randomly selected areas and 0–3 and ≥ 4 in the most mitotically active areas).

Although there were few deaths, we analyzed 5-year OS rates. For specimens in which the number of mitoses was counted in a randomly selected 10-HPF area, the 5-year OS rates with 0–1 mitoses/10 HPF and 2–10 mitoses/10 HPF were 100% and 96.6%, respectively ($P = 0.0033$; Fig. 3D). For specimens in which the number of mitoses in the most mitotically active 10-HPF area was counted, the 5-year OS rates for specimens with 0–1 mitoses/10 HPF, 2–10 mitoses/10 HPF, and > 11 mitoses/10 HPF were 100%, 95.5%, and 82.1%, respectively ($P = 0.1471$; Fig. 3E). For specimens in which the overall mean number of mitoses was calculated, the 5-year OS rates for specimens with an overall mean mitosis < 0.2 and a mean mitosis ≥ 0.2 were 98.7% and 0%, respectively ($P < 0.0001$; Fig. 3F).

3.7. New recommended definition for AC

Based on these results, we recommend that the definition for AC include comedo-like necrosis and/or an overall mean number of mitosis of ≥ 0.2 . Based on this definition, we observed AC in 7 cases rather than the 12 cases specified in the records. The 5-year RFS rates for patients with TC and AC by this new definition were 97.2% and 0%, respectively ($P < 0.0001$; Fig. 4). We observed lymph node metastasis in 14 patients with TC and 3 patients with AC; however, there was no significant difference between these two rates ($P = 0.166$).

4. Discussion

Our results favor a new histological definition for AC. In this study, we found that the overall mean number of mitoses was superior to calculate the number of mitoses in a randomly selected 10-HPF area or in the most mitotically active 10-HPF area. Furthermore, the presence of comedo-like necrosis was significantly associated with tumor recurrence and patient death when compared to the presence of coagulative necrosis.

A high mitotic count is generally reported as an unfavorable factor in patients with tumors. In general, selecting the most mitotically active 10-HPF area is the preferred method for calculating mitotic counts [13–15]. Most of the articles on pulmonary neuroendocrine tumors, including the first publication on the definition of LCNEC, counted the mitoses in the most mitotically active areas [7,16,17]. Our results indicated that the mitotic counts in both ACs and TCs were distributed heterogeneously. Surprisingly, 21 of 26 cases with a mitotic count exceeding 2 mitoses in only 1 HPF showed no tumor recurrence or patient death. When compared with a randomly selected 10-HPF area or the most mitotically active 10-HPF area, ≥ 2 mitoses in only 1 HPF had the least selection bias to go beyond the diagnostic criteria for TC. Furthermore, when we selected the most mitotically active areas, 20 cases went

Table 1
Correlation between clinicopathological factors and unfavorable outcomes.

	Recurrence/death		P-value ^a	Lymph node metastasis		P-value ^b
	(–)	(+)		(–)	(+)	
Original diagnosis						
Typical	63	5	<0.0001	55	12	0.1192
Atypical	7	5		7	5	
Age (years)	52.5	67.1	0.002	54.1	54.4	1.00
Gender						
Female	45	4	0.1944	38	10	0.8536
Male	25	6		24	7	
Symptom						
None	43	7	0.8271	38	11	0.5524
Present	24	3		23	4	
Smoking						
Never	31	1	0.0767	27	4	0.5267
Former	25	7		24	8	
Present	10	2		10	2	
Location						
Central	44	5	0.3645	38	11	0.7971
Non-central	26	5		24	6	
Tumor size (cm)	2.92	3.23	0.936	3.05	2.69	0.9952
Lymph node metastasis						
Absent	55	7	0.4563	–	–	–
Present	14	3		–	–	
Pathologic stage						
I and II	61	6	0.0029	–	–	–
III and IV	8	4		–	–	
Infiltrative growth						
Absent	29	4	0.4805	27	5	0.4051
Present	41	6		35	12	
Blood-vessel invasion						
Absent	46	3	0.088	39	6	0.0417
Present	27	7		23	11	
Lymph-vessel invasion						
Absent	46	5	0.2246	45	5	0.0017
Present	24	5		17	12	
Pleural invasion						
Absent	70	9	<0.0001	61	17	1.00
Present	0	1		1	0	
Stroma						
Capillary rich	52	5	0.1467	43	13	0.7649
Fibrous rich	18	5		19	4	
Active fibroblastic proliferation						
Absent	40	2	0.0184	34	7	0.3179
Present	30	8		28	10	
Chromatin pattern						
Fine	40	7	0.5614	40	6	0.0304
Coarse	30	3		22	11	
Nucleolus						
Inconspicuous	57	8	0.7714	54	10	0.0085
Conspicuous	13	2		8	7	
Nuclear pleomorphism						
Absent	59	8	0.5684	56	10	0.0019
Present	11	2		6	7	

^a Log-rank test.^b Fisher's exact test.

beyond the diagnostic criteria for AC. These results indicated that the latest WHO criteria for diagnosing ACs is strict only evaluating in 10 HPFs.

We also attempted to determine if there was a pattern to the distribution of mitoses (such as within the tumor's center or in the periphery); however, we did not see a uniform tendency (unpublished results). Travis et al. counted the number of mitoses in

pulmonary neuroendocrine tumors in three sets of 10 HPF and calculated the mean [6,16]. Their results indicated that one set of 10 HPFs is not enough to evaluate the mitotic activity because of the mitotic heterogeneity in carcinoid tumors. Tiny material, such as biopsy specimens may not be suitable for a final diagnosis of TC or AC, because there may not be enough HPFs. The results of our study are based on surgically resected materials,

Table 2
Correlation between necrosis and unfavorable outcomes.

	Recurrence/death		<i>P</i> -value ^a	Lymph node metastasis		<i>P</i> -value ^b
	(–)	(+)		(–)	(+)	
Comedo-like necrosis						
Absent	69	6	<0.0001	60	14	0.0639
Present	1	4		2	3	
Coagulative necrosis						
Absent	62	8	0.3143	55	14	0.4414
Present	8	2		7	3	
Pyknotic change						
Absent	25	1	0.0954	22	3	0.2405
Present	45	9		40	14	
Cholesterol cleft						
Absent	67	9	0.4762	59	16	1.00
Present	3	1		3	1	
Hemorrhage						
Absent	47	4	0.102	41	9	0.3176
Present	23	6		21	8	
Psammomatous body						
Absent	61	6	0.0559	53	13	0.4611
Present	9	4		9	4	

^a Log-rank test.
^b Fisher's exact test.

and it is unclear whether calculating the overall mean number of mitoses would be applicable for materials acquired from a biopsy.

The presence of necrosis is another criterion for differentiating TC from AC. In the current study, we analyzed the necrosis separating comedo-like and coagulative necrosis. We found that comedo-like necrosis was significantly associated with tumor recurrence and patient death when compared to coagulative necrosis. In fact, we did not see tumor recurrence or patient death in the cases with only coagulative necrosis. This may be because coagulative necrosis is not related to the tumor's malignant behavior but rather it may be related to preoperative biopsy or erosion. Travis et al. also pointed out that the large zone of infarct-like necrosis (corresponding to current coagulative necrosis) was not seen in ACs but was seen in high-grade neuroendocrine carcinomas [16].

It has been reported that hemorrhage, psammoma bodies, and/or cholesterol cleft form after the degeneration and disappear-

ance of tumor cells [18–20]. However, we found that these factors were not associated with the patients' clinical outcomes in our study. We observed pyknotic cells, which can sometimes be confused with incipient necrosis, in 69% of patients with both ACs and TCs. Similar to previous reports, our results demonstrated that the presence of pyknotic cells was not associated with the patients' clinical outcomes [7,8].

The presence of an active fibroblastic proliferation is one of the known unfavorable clinical outcomes in patients with pulmonary adenocarcinomas [12]. In our study, we observed active fibroblastic proliferation in 48% of the cases, which was significantly associated with tumor recurrence or patient death (*P*=0.0184). All patients who had tumor specimens with active fibroblastic proliferation had undergone a preoperative biopsy and/or cytologic diagnostic procedure. However, these results did not exclude the possibility of the active fibroblastic proliferation induced by preoperative diagnostic procedures. Therefore, it will be necessary to study a larger number of patients who have not undergone

Table 3
Correlation between three mitotic count methods and unfavorable outcomes.

	Recurrence/death		<i>P</i> -value ^a	Lymph node metastasis		<i>P</i> -value ^b
	(–)	(+)		(–)	(+)	
Randomly selected 10 HPF						
<2	65	6	0.003	56	14	0.3957
≥2	5	4		6	3	
<2	69	8	0.002	60	16	0.522
≥2	1	2		2	1	
Highest 10 area						
<2	21	1	0.0575	21	1	0.0641
2–10	35	4		28	11	
>11	14	5	0.0238	13	5	0.180
0–3	37	2		28	11	
≥4	33	8		34	6	
Mean mitosis						
<0.2	69	6	<0.0001	60	14	0.0639
≥0.2	1	4		2	3	

^a Log-rank test.
^b Fisher's exact test.

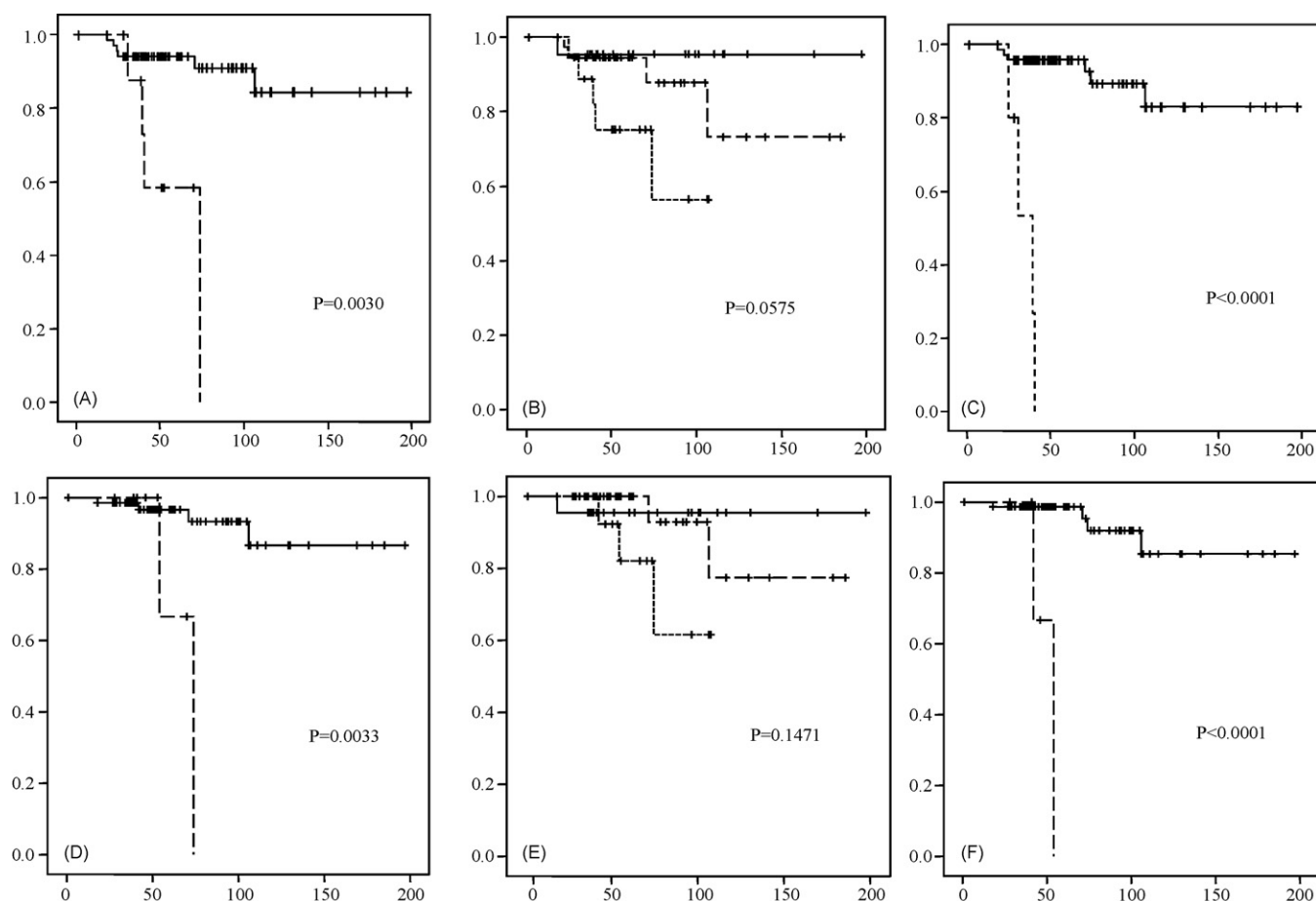


Fig. 3. The recurrence-free survival (RFS) rates for patients with ACs. (A) For cases where the number of mitoses was calculated in randomly selected 10-high-power field (HPF) areas, the 5-year RFS rates for cases with 0–1 mitoses/10 HPF (solid line) and 2–10 mitoses/10 HPF (dashed line) were 94.2% and 58.3%, respectively. (B) For cases where the number of mitoses was calculated in the most mitotically active 10-HPF area, the 5-year RFS rates for cases with 0–1 mitoses/10 HPF (solid line), 2–10 mitoses/10 HPF (dashed line), and ≥ 11 mitoses/10 HPF (square-dot line) were 95.5%, 94.6% and 75.2%, respectively. (C) For cases where the average mitotic count was calculated, the 5-year RFS rates for cases with <0.2 mitoses/10 HPF (solid line) and ≥ 0.2 mitoses/10 HPF (dashed line) were 95.9% and 0%, respectively. (D) specimens with randomly selected 10-HPF are, the 5-year OS rates with 0–1 mitoses/10 HPF and 2–10 mitoses/10 HPF were 100% and 96% respectively. (E) Specimens with more mitotically active 10 HPF area, 5-year OS rates with 0–1 mitoses/10 HPF, 2–10 mitoses/10 HPF, and >11 mitoses/10 HPF were 100%, 95%, and 82% respectively. (F) Specimens with overall mean of mitoses, the 5-year OS rates with mean <0.2 and a mean mitosis >0.2 were 98.7% and 0% respectively.

preoperative diagnostic procedures to determine the significance of active fibroblastic proliferation in patients with carcinoid tumors.

Although our cohort was not large enough to analyze each pathologic disease stage, patients with p-stage III or IV disease were more likely to experience a tumor recurrence or death than patients with p-stage I or II disease ($P=0.00294$). Adapting the TNM classification scheme for patients with lung carcinoid tumors is an important factor for predicting survival [21]. However, one of the important factors of TNM classification – lymph node metastasis – was not correlated with tumor recurrence or patient death in our study. Currently, there are conflicting reports about the effect of lymph node metastasis for survival in patients with carcinoid tumors. Some studies have found that in patients with TC, lymph node metastasis did not affect survival [8,22]. However, if some histologic factors can predict the risk of lymph node metastasis, they may be useful for determining the extent of the regional lymph node dissection for surgeons. In our study, the presence of blood-vessel invasion, lymph-vessel invasion, coarse chromatin pattern, nucleolus prominence, and nuclear pleomorphism were correlated with lymph node metastasis, but mitotic activity and/or the presence of necrosis were not. Based on these factors, the presence of a coarse chromatin pattern,

nucleolus prominence, and nuclear pleomorphism can be diagnostic features for lymph node metastasis in preoperative biopsy materials.

We found that counting the overall mean number of mitoses was superior to the other methods for determining the specimens' mitotic counts. However, counting the number of mitoses in all fields in each case at a high magnification may be unsuitable for daily practice because of the substantial time involved. Therefore, based on our results, we made a flow chart (Fig. 5) to aid in the diagnosis of low- and intermediate-grade neuroendocrine carcinomas.

In short, we concluded that the mitotic figures of carcinoid tumors have a heterogeneous distribution. To reduce the selection biases of fields and observer variation, we recommended calculating the overall mean number of mitoses. We also pointed out the potential for overdiagnosing carcinoid tumors using the method for counting mitoses in the most mitotically active 10-HPF. Furthermore, we concluded that comedo-like necrosis was a more important pattern to look for than coagulative necrosis for predicting tumor recurrence or patient death. The combination of calculating the mean number of mitoses and the presence of comedo-like necrosis should be added to the current criteria for diagnosing ACs.

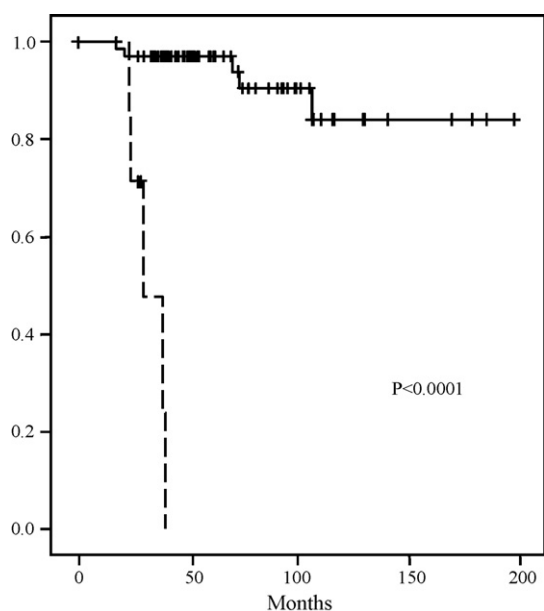


Fig. 4. Based on our recommended criteria for atypical carcinoids, the 5-year disease-free survival rates for patients with typical carcinoids (solid line) and atypical carcinoids (comedo-like necrosis and/or an average mitosis ≥ 0.2 mitoses/10 HPF; dashed line) were 97.2% and 0%, respectively.

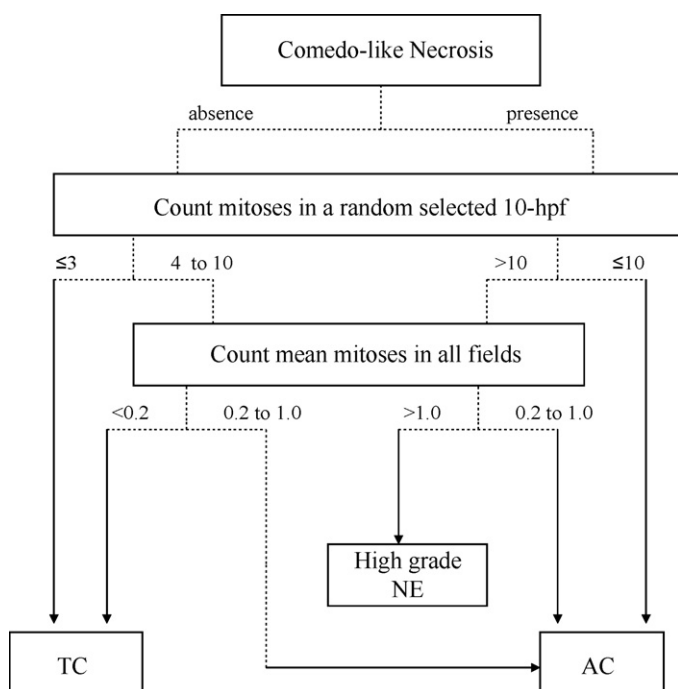


Fig. 5. A dichotomous tree for diagnosing neuroendocrine tumors of the lung, according to our results. Abbreviations: TC, typical carcinoid; AC, atypical carcinoid; NE, neuroendocrine tumor; HPF, high-power field.

Conflict of interest statement

There is no conflict of interest with regard to this manuscript.

Acknowledgements

We would like to thank Susan Cweren for her skillful technical assistance and Alyson Todd for her skillful English editing.

This work was supported in part by the grant from the US Department of Defense PROSPECT W81XWH-07-1-0306.

References

- [1] Fink G, Krelbaum T, Yellin A, Bendayan D, Saute M, Glazer M, et al. Pulmonary carcinoid: presentation, diagnosis, and outcome in 142 cases in Israel and review of 640 cases from the literature. *Chest* 2001;119:1647–51.
- [2] Hage R, de la Riviere AB, Seldenrijk CA, van den Bosch JM. Update in pulmonary carcinoid tumors: a review article. *Ann Surg Oncol* 2003;10:697–704.
- [3] Granberg D, Wilander E, Oberg K, Skogseid B. Prognostic markers in patients with typical bronchial carcinoid tumors. *J Clin Endocrinol Metab* 2000;85:3425–30.
- [4] Cooper WA, Thourani VH, Gal AA, Lee RB, Mansour KA, Miller JL. The surgical spectrum of pulmonary neuroendocrine neoplasms. *Chest* 2001;119:14–8.
- [5] Garcia-Yuste M, Matilla JM, Cueto A, Paniagua JM, Ramos G, Cañizares MA. Typical and atypical carcinoid tumours: analysis of the experience of the Spanish Multi-centric Study of Neuroendocrine Tumours of the Lung. *Eur J Cardiothorac Surg* 2007;31:192–7.
- [6] Travis WD, Rush W, Flieder DB, Falk R, Fleming MV, Gal AA. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. *Am J Surg Pathol* 1998;22:934–44.
- [7] Beasley MB, Thunnissen FB, Brambilla E, Hasleton P, Steele R, Hammar SP, et al. Pulmonary atypical carcinoid: predictors of survival in 106 cases. *Hum Pathol* 2000;31:1255–65.
- [8] Thomas Jr CF, Tazelaar HD, Jett JR. Typical and atypical pulmonary carcinoids: outcome in patients presenting with regional lymph node involvement. *Chest* 2001;119:1143–50.
- [9] Asamura H, Kameya T, Matsuno Y, Noguchi M, Tada H, Ishikawa Y, et al. Neuroendocrine neoplasms of the lung: a prognostic spectrum. *J Clin Oncol* 2006;24:70–6.
- [10] Beasley MB, Thunnissen FB, Hasleton PhS, Barbareschi M, Pugatch B. Carcinoid tumour. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. *Pathology and genetics: tumors of the lung, pleura, thymus and heart*. Lyon, France: IARC; 2004. p. 59–62.
- [11] Arrigoni MG, Woolner LB, Bernatz PE. Atypical carcinoid tumors of the lung. *J Thorac Cardiovasc Surg* 1972;64:413–21.
- [12] Noguchi M, Morikawa A, Kawasaki M, Matsuno Y, Yamada T, Hirohashi S, et al. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. *Cancer* 1995;75:2844–52.
- [13] Kempson RL. Mitosis counting-II (editorial). *Hum Pathol* 1976;7:482–3.
- [14] Baak JPA. Mitosis counting in tumors. *Hum Pathol* 1990;21:683–5.
- [15] van Diest PJ, Baak JP, Matze-Cok P, Wisse-Brekelmans EC, van Galen CM, Kurver PH, et al. Reproducibility of mitosis counting in 2,469 breast cancer specimens: results from the Multicenter Morphometric Mammary Carcinoma Project. *Hum Pathol* 1992;23:603–7.
- [16] Travis WD, Linnoila RI, Tsokos MG, Hitchcock CL, Cutler Jr GB, Nieman L, et al. Neuroendocrine tumors of the lung with proposed criteria for large-cell neuroendocrine carcinoma: an ultrastructural, immunohistochemical, and flow cytometric study of 35 cases. *Am J Surg Pathol* 1991;15:529–53.
- [17] Ruge M, Fassan M, Clemente R, Rizzardi G, Giacomelli L, Pennelli G, et al. Bronchopulmonary carcinoid: phenotype and long-term outcome in a single-institution series of Italian patients. *Clin Cancer Res* 2008;14:149–54.
- [18] Shimamoto Y, Suzuki A, Hashimoto T, Nishiwaki Y, Kodama T, Yoneyama T, et al. Prognostic implications of fibrotic focus (scar) in small peripheral lung cancers. *Am J Surg Pathol* 1980;4:365–73.
- [19] McCluggage WG, Lyness RW, Atkinson RJ, Dobbs SP, Harley I, McClelland HR, et al. Morphological effects of chemotherapy on ovarian carcinoma. *J Clin Pathol* 2002;55:27–31.
- [20] Liu-Jarin X, Stoopler MB, Raftopoulos H, Ginsburg M, Gorenstein L, Borczuk AC. Histologic assessment of non-small cell lung carcinoma after neoadjuvant therapy. *Mod Pathol* 2003;160:1102–8.
- [21] Travis WD, Giroux DJ, Chansky K, Crowley J, Asamura H, Brambilla E, et al. The IASLC Lung Cancer Staging Project: proposals for the inclusion of bronchopulmonary carcinoid tumors in the forthcoming (seventh) edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* 2008;3:1213–23.
- [22] Schreurs AJ, Westermann CJ, van den Bosch JM, Vanderschueren RG, Brutel de la Riviere A, Knaepen PJ. A twenty-five-year follow-up of ninety-three resected typical carcinoid tumors of the lung. *Thorac Cardiovasc Surg* 1992;104:1470–5.

ORIGINAL ARTICLE

Epidermal growth factor receptor regulates MET levels and invasiveness through hypoxia-inducible factor-1 α in non-small cell lung cancer cells

L Xu^{1,8}, MB Nilsson^{1,8}, P Saintigny¹, T Cascone¹, MH Herynk¹, Z Du², PG Nikolinakos¹, Y Yang¹, L Prudkin³, D Liu⁴, JJ Lee⁴, FM Johnson¹, K-K Wong⁵, L Girard⁶, AF Gazdar⁶, JD Minna^{6,7}, JM Kurie¹, II Wistuba^{2,3} and JV Heymach^{1,2}

¹Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ²Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ³Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ⁴Department of Biostatistics and Applied Mathematics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ⁵Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; ⁶Department of Pharmacology, Hamon Center for Therapeutic Oncology Research, Simmons Comprehensive Cancer Center, The University of Texas Southwestern Medical Center, Dallas, TX, USA and ⁷Department of Internal Medicine, The University of Texas Southwestern Medical Center, Dallas, TX, USA

Recent studies have established that amplification of the *MET* proto-oncogene can cause resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) in non-small cell lung cancer (NSCLC) cell lines with EGFR-activating mutations. The role of non-amplified *MET* in EGFR-dependent signaling before TKI resistance, however, is not well understood. Using NSCLC cell lines and transgenic models, we demonstrate here that EGFR activation by either mutation or ligand binding increases *MET* gene expression and protein levels. Our analysis of 202 NSCLC patient specimens was consistent with these observations: levels of *MET* were significantly higher in NSCLC with *EGFR* mutations than in NSCLC with wild-type *EGFR*. EGFR regulation of *MET* levels in cell lines occurred through the hypoxia-inducible factor (HIF)-1 α pathway in a hypoxia-independent manner. This regulation was lost, however, after *MET* gene amplification or overexpression of a constitutively active form of HIF-1 α . EGFR- and hypoxia-induced invasiveness of NSCLC cells, but not cell survival, were found to be *MET* dependent. These findings establish that, absent *MET* amplification, EGFR signaling can regulate *MET* levels through HIF-1 α and that *MET* is a key downstream mediator of EGFR-induced invasiveness in EGFR-dependent NSCLC cells.

Oncogene advance online publication, 15 February 2010; doi:10.1038/onc.2010.16

Keywords: EGFR; *MET*; non-small cell lung cancer; HIF-1 α ; invasiveness

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer death in the United States. Epidermal growth factor receptor (EGFR)-activating mutations have been described in a subset of NSCLC patients, and activated EGFR is known to influence tumor cell survival, proliferation, angiogenesis, and invasiveness (Lynch *et al.*, 2004; Paez *et al.*, 2004; Pao *et al.*, 2004; Janne *et al.*, 2005; Pao and Miller, 2005; Ciardiello and Tortora, 2008). EGFR tyrosine kinase inhibitors (TKIs) such as erlotinib and gefitinib are clinically active in 10–20% of NSCLC patients (Fukuoka *et al.*, 2003; Kris *et al.*, 2003; Shepherd *et al.*, 2005; Thatcher *et al.*, 2005). Activating mutations within the *EGFR* tyrosine kinase domain including an amino acid substitution at exon 21 (L858R) and in-frame deletions in exon 19 were found to be predictors of clinical response to EGFR TKIs (Lynch *et al.*, 2004; Paez *et al.*, 2004; Pao *et al.*, 2004).

Recent evidence suggests that in NSCLC cells activating EGFR mutations or amplification of the *MET* proto-oncogene caused acquired resistance to EGFR TKIs by driving activation of the PI3K pathway (Engelman *et al.*, 2007). The role of *MET* in EGFR-dependent signaling before the emergence of TKI resistance is not well understood; however, *MET* is regulated by hypoxia and hypoxia-inducible factor-1 α (HIF-1 α) and is thought to contribute to invasive tumor growth (Pennacchietti *et al.*, 2003). The *MET* protein is a receptor tyrosine kinases whose activation can cause malignant transformation and tumorigenesis (Cooper *et al.*, 1986; Park *et al.*, 1987; Stabile *et al.*, 2004). Upon ligand binding, *MET* activates downstream signaling molecules including PI3K, Src, and signal transducer and activator of transcription-3 (Rosario and Birchmeier, 2003), triggering the key metastatic steps of cell dissociation (Qiao *et al.*, 2002), migration (Yi *et al.*, 1998), and invasion (Bredin *et al.*, 2003). *MET* is overexpressed in multiple malignancies and is associated with aggressive disease (Peruzzi and Bottaro, 2006). In

Correspondence: Dr JV Heymach, Departments of Thoracic/Head and Neck Medical Oncology and Cancer Biology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Unit 432, Houston, TX 77030, USA.

E-mail: jheyman@mdanderson.org

⁸These authors contributed equally to this work.

Received 22 August 2009; revised 7 January 2010; accepted 11 January 2010

NSCLC, MET levels are elevated in resected tumors compared with normal tissue (Liu and Tsao, 1993; Ichimura *et al.*, 1996; Olivero *et al.*, 1996), and high expression of the MET ligand hepatocyte growth factor (HGF) is associated with aggressive disease and a poor prognosis (Siegfried *et al.*, 1998).

Recent studies have suggested a link between EGFR signaling and MET. Expression and phosphorylation of EGFR and MET correlate in multiple malignancies (Weinberger *et al.*, 2005). Aberrant EGFR activation results in elevated MET phosphorylation in thyroid carcinoma cells (Bergstrom *et al.*, 2000). EGFR function has been implicated in HGF-induced hepatocyte proliferation (Scheving *et al.*, 2002) and is required for MET-mediated colon cancer cell invasiveness (Pai *et al.*, 2003). Recent studies of phosphoprotein networks reveal an association between EGFR and MET activation (Huang *et al.*, 2007; Guo *et al.*, 2008), and have reported direct crosstalk between EGFR and the MET (Jo *et al.*, 2000; Huang *et al.*, 2007).

A plausible link between the EGFR and MET pathways is HIF-1, which has two subunits (HIF-1 α and HIF-1 β), and is known to contribute to tumor cell motility and invasiveness. EGF has been shown to modulate HIF-1 α levels in prostate, breast, and lung cancer cell lines (Zhong *et al.*, 2000; Phillips *et al.*, 2005; Peng *et al.*, 2006), and positive correlations between EGFR and HIF-1 α expression have been observed in NSCLC (Hirami *et al.*, 2004; Swinson *et al.*, 2004).

Here, we have used clinical specimens, transgenic mouse models, and cell lines to investigate the hypothesis that EGFR signaling may regulate MET levels through HIF-1 α but that *MET* amplification, which occurs in EGFR TKI resistance, would uncouple MET levels from EGFR regulation. We hypothesized further that EGFR-induced invasiveness, like hypoxia-induced invasiveness, is mediated downstream at least in part by the HIF-1 α /MET axis.

Results

EGFR-activating mutations are associated with elevated levels of MET in NSCLC clinical samples

To investigate a possible association between EGFR activation and MET in clinical specimens, we evaluated MET levels by immunohistochemistry and assessed *EGFR* mutations in 202 human NSCLC clinical specimens. Out of 202 samples, 22 had detectable *EGFR* mutations. Specimens were immunostained for MET and scored based on an intensity score (0, 1, 2, or 3) and an extension percentage. The final score was the product of these two values. The mean score for MET expression was 39.46 ± 64.52 . Therefore, a score of 40 was considered the cutoff for classifying low and high levels of MET expression. The mean MET expression score was significantly higher in specimens with mutated *EGFR* (73.64 ± 70.68) than in specimens with WT *EGFR* (48.72 ± 71.72 ; $P=0.04$; Figure 1a). Furthermore, 37% of NSCLC tumors with WT *EGFR*

expressed high levels of membranous MET, whereas 68% of NSCLC tumors with mutated *EGFR* expressed high levels of membranous MET ($P=0.005$; Figure 1b). Among adenocarcinomas with EGFR-activating mutations, we did not observe any association between EGFR expression and survival. However, considering the small sample size, no definitive conclusions can be drawn.

EGFR activation modulates MET expression in transgenic murine models of NSCLC

We investigated whether a similar association between EGFR-activating mutations and MET expression occurred in murine models of NSCLC. We used transgenic mice with lung tumors driven by lung-specific mutated K-RAS or activating EGFR mutation (Forsythe *et al.*, 1996; Johnson *et al.*, 2001). Lung tumor sections were immunostained for MET and scored as described above. *K-RAS*-driven lesions had an average score of 6.75, whereas tumors with *EGFR*-activating mutations had an average staining score of 40.65 (Figure 1c; $P<0.001$). Treatment of mice bearing EGFR-driven lung tumors with the EGFR TKI erlotinib (50 mg/kg/day) for 48 h abolished MET, providing evidence that MET levels were regulated by EGFR activation.

EGFR-activating mutations are associated with elevated HIF-1 α and MET levels in NSCLC cell lines

Given our finding that tumors with *EGFR* mutations exhibit higher MET expression, we investigated MET regulation by EGFR and its role in EGFR-mediated NSCLC invasiveness. We evaluated *MET* RNA levels in NSCLC cell lines by performing gene expression analysis on gene arrays of 53 previously characterized NSCLC lines (eight lines with mutated *EGFR*) (GEO 4824) (Zhou *et al.*, 2006). *MET* RNA levels were significantly higher in *EGFR*-mutated cell lines than in NSCLC cell lines expressing WT *EGFR* (Figure 2a; $P=0.002$); however, *MET* expression levels in cell lines with *K-RAS* mutations were not significantly different compared with cell lines with WT *K-RAS*. Moreover, we observed a significant association between *EGFR* gene copy number (>4 copies using RT-PCR) and levels of *MET* expression ($P=0.03$, Figure 2b).

We evaluated MET protein levels in NSCLC with or without *EGFR*-activating mutations and observed constitutive EGFR phosphorylation in cell lines with mutated *EGFR*, which was associated with increased phosphorylated MET (p-MET) and MET expression (Figure 2c). Cell lines with *EGFR*-activating mutations were positive for HIF-1 α expression in normoxia. HCC827 cells, which exhibited the most robust expression of p-EGFR, produced the highest levels of HIF-1 α , p-MET, and MET. Western data are supported by ELISA analysis showing higher levels of p-EGFR, p-MET, and HIF-1 α in cell lines with EGFR-activating mutations compared with cells with WT *EGFR* (Figures 2d–f).

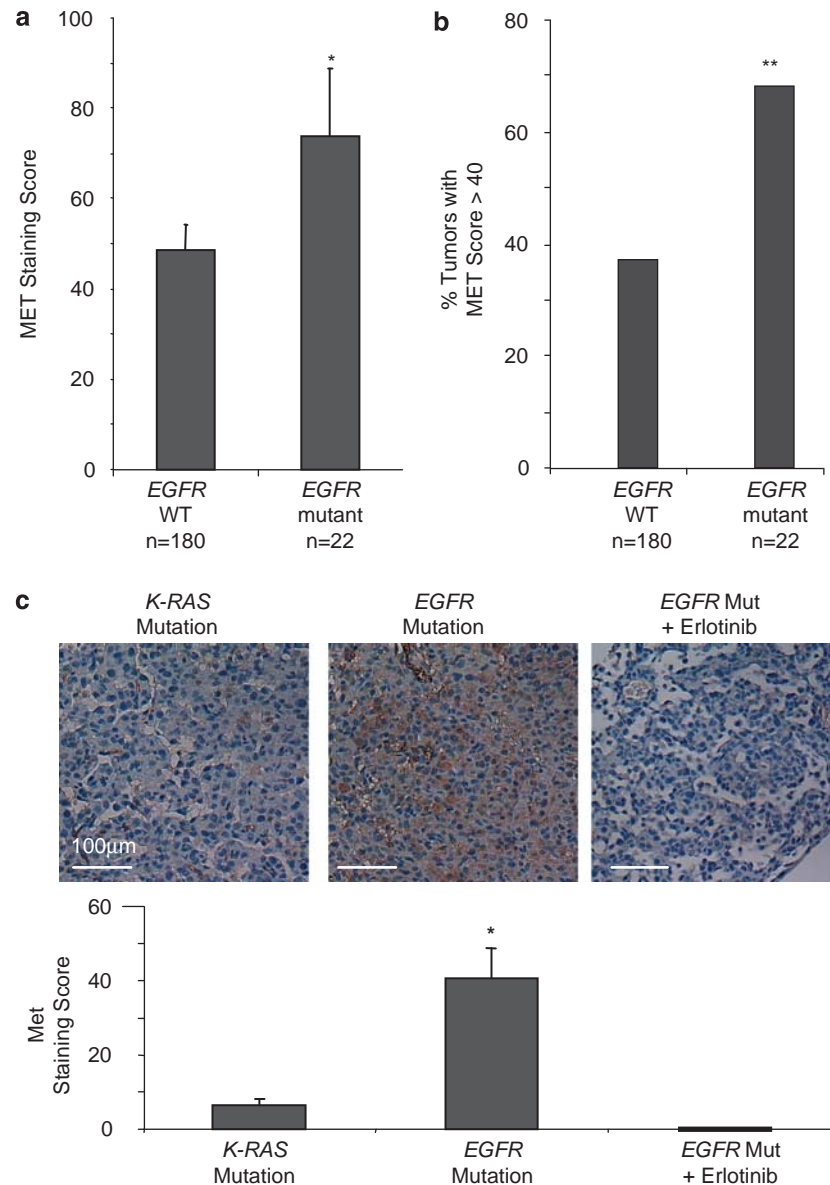


Figure 1 Elevated MET and HGF expression correlates with *EGFR*-activating mutations in NSCLC tumor samples. NSCLC clinical specimens ($n = 202$) were immunostained with anti-MET ab and scored (a). *EGFR*-activating mutations correlated with elevated levels of MET. Bars, s.e.m.; * $P < 0.05$. (b) Data are presented as the percentage of tumors with high MET expression; ** $P < 0.005$. (c). Murine lung tumors driven by *K-RAS* or *EGFR*-activating mutations were immunostained with anti-MET ab, and positive staining was quantified. Weak or negative MET staining was observed in *K-RAS*-driven tumors, whereas tumors with *EGFR*-activating mutations exhibited elevated MET expression. Erlotinib treatment diminished MET expression. Representative images are shown. Columns, mean score; bars, s.e.m. * $P < 0.001$.

Activated EGFR modulates p-MET, MET, and HIF-1 α
We treated HCC827 cells with 1 μ M of erlotinib for 12 h and evaluated p-MET, MET, and HIF-1 α levels. Erlotinib reduced p-MET and MET protein (Figure 3a). EGFR inhibition resulted in diminished HIF-1 α levels. p-MET, MET, and p-EGFR were further analyzed by ELISA assay (Figure 3b). Consistent with data obtained by western blot, erlotinib decreased p-EGFR ($P = 0.009$), p-MET ($P = 0.1$), and MET ($P = 0.001$) levels. As HIF-1 α is known to regulate *MET* transcription, we determined whether mutated *EGFR* would regulate *MET* mRNA levels. We treated HCC827 cells

with or without erlotinib (1 μ M) for 12 h and collected RNA for RT-PCR to evaluate changes in *MET* mRNA relative to *GAPDH* RNA. Inhibition of EGFR activity resulted in approximately a 50% decrease in *MET* RNA compared with control levels (Figure 3c).

To further show that EGFR signaling modulates HIF-1 α and MET protein expression, we transfected HCC827 cells with control siRNA and EGFR-, HIF-1 α -, and MET-targeting siRNA. Knockdown of EGFR decreased p-MET, MET, and HIF-1 α levels. HIF-1 α -targeting siRNA did not alter EGFR expression but reduced MET expression and activation, whereas MET

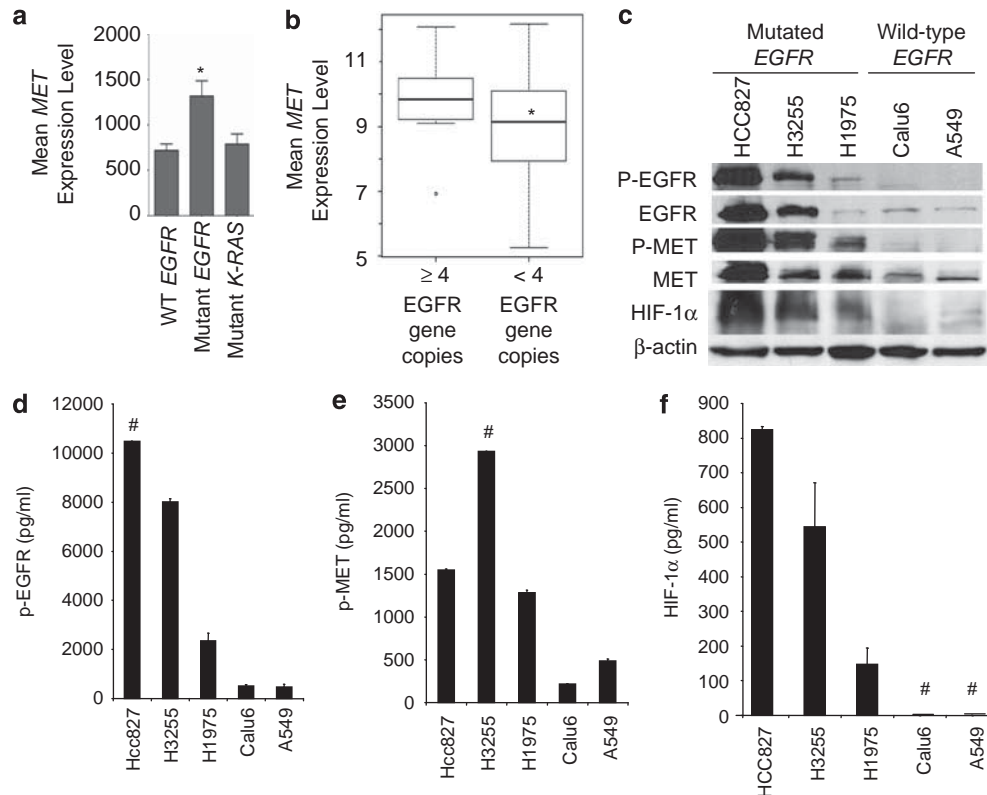


Figure 2 EGFR-activating mutations are associated with elevated MET and HIF-1 α levels in NSCLC cell lines. (a) Gene expression analysis was performed on gene arrays of 53 NSCLC lines. MET expression was elevated in NSCLC cell lines harboring EGFR-activating mutations; * $P=0.002$. (b) MET expression in 53 NSCLC cell lines with high EGFR gene copy number (>4 copies) vs low copy number (<4 copies); * $P=0.03$. (c) Western blot was used to evaluate pEGFR, EGFR, p-MET, MET, and HIF-1 α expression in NSCLC cell lines expressing WT EGFR or mutationally activated EGFR. The presence of EGFR-activating mutations was associated with increased levels of p-MET, MET, and HIF-1 α . (d–f) ELISA assay was used to analyze levels of p-EGFR (d), p-MET (e), and HIF-1 α (f) in NSCLC cell lines expressing WT EGFR or mutationally activated EGFR. # indicates samples that were out of range.

siRNA reduced MET but not EGFR or HIF-1 α levels (Figure 3d), indicating that HIF-1 α and MET are downstream of EGFR. Similar results were obtained by HIF-1 α ELISA assay. siRNA directed against EGFR but not MET decreased HIF-1 α levels ($P=0.009$; Figure 3e).

MET amplification has been described in a subset of NSCLC patients (Zhao *et al.*, 2005; Engelman *et al.*, 2007). To determine whether MET amplification would result in MET expression that was independent of EGFR, we treated H1993 NSCLC cells, which harbor an amplified MET allele (Engelman *et al.*, 2007; Lutterbach *et al.*, 2007), with erlotinib, and evaluated p-EGFR, EGFR, p-MET and MET levels. In contrast to the EGFR-dependent cell lines tested, pharmacological inhibition of EGFR did not diminish MET expression in this cell line (Figure 3f).

Previous studies suggested that activated EGFR can directly induce phosphorylation of MET (Bergstrom *et al.*, 2000; Jo *et al.*, 2000). To evaluate the effect of EGFR activation on MET in NSCLC, we stimulated A549 cells with EGF with or without erlotinib. Phosphorylated EGFR was detected 30 min after ligand

stimulation, and EGFR activation was inhibited with erlotinib (Figure 3g). EGFR levels decreased 12 h after the addition of EGF, which may have been a result of receptor internalization. EGF stimulation triggered rises in p-MET levels at 30 min, suggesting that EGFR directly activated MET. p-MET levels remained detectable 6 h after EGF stimulation. Prolonged exposure (24 h) to EGF resulted in increased levels of MET protein (Figure 3h).

EGFR-mediated invasion of NSCLC cells is MET dependent

To show that MET activation increases invasiveness in NSCLC and that this can be abrogated with the MET TKI, PHA-665752, we treated A549 and HCC827 tumor cells with the MET ligand HGF alone or with PHA-665752. Cell invasion was measured using Matrigel-coated Boyden chambers. In both cell lines, HGF stimulation resulted in a significant increase in invasiveness, and this was inhibited with the addition of PHA-665752 ($P<0.05$; Figure 4a). As EGFR activation has been shown to modulate tumor cell invasion in multiple

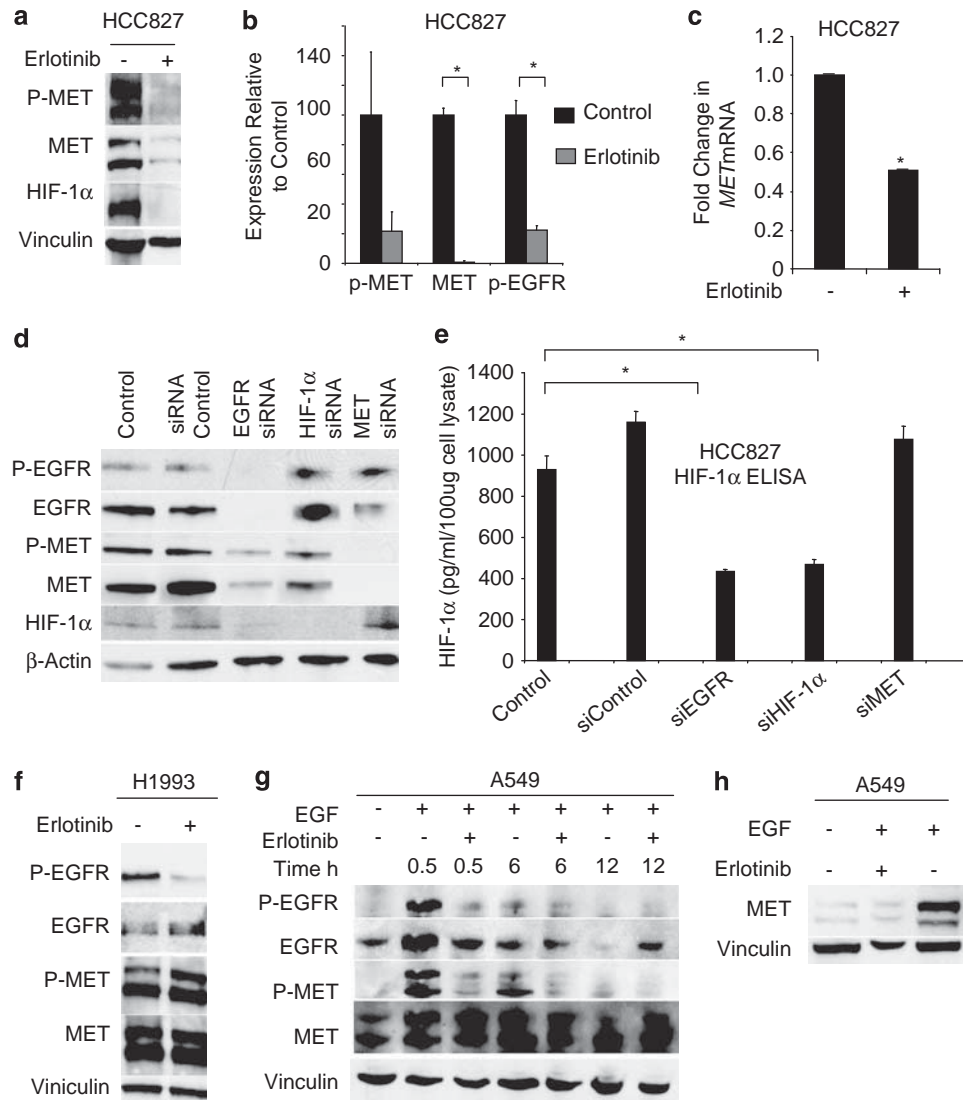


Figure 3 EGFR activation modulates levels of MET and p-MET in NSCLC cell lines. (a) HCC827 cells were treated with 1 μ M erlotinib for 12 h. Protein lysates were collected after 8 h, and p-MET, MET, and HIF-1 α levels were evaluated by western blot. (b) HCC827 cells were treated with 1 μ M erlotinib for 12 h, and levels of p-MET, MET, and p-EGFR were evaluated by ELISA assay. * P <0.01. (c) HCC827 cells were treated with 1 μ M erlotinib for 12 h, and *MET* mRNA levels were evaluated by RT-PCR. Inhibition of EGFR activation decreased *MET* RNA. Bars, s.d.; * P <0.001. (d) HCC827 cells were transfected with siRNA oligonucleotides directed against EGFR, HIF-1 α , MET, and non-targeting control siRNA. After 72 h, protein lysates were collected and western blot was performed. (e) HIF-1 α levels are decreased in HCC827 cells transfected with siRNA-targeting EGFR and HIF-1 α as determined by ELISA. * P <0.05. (f) H1973 cells, which have an amplified *MET* allele, were treated with 1 μ M erlotinib for 12 h, and p-EGFR, EGFR, p-MET and MET expression were evaluated by immunoblot. (g) A549 cells were serum starved for 12 h and treated with EGF (60 ng/ml) with or without 1 μ M erlotinib. Protein lysates were collected at the indicated times, and EGFR and MET activation were evaluated by immunoblot. (h) A549 cells were serum starved for 12 h and then stimulated with 60 ng/ml EGF with or without 1 μ M erlotinib for 24 h.

cell types including NSCLC (Hamada *et al.*, 1995; Damstrup *et al.*, 1998), we investigated whether EGFR activation's effect on tumor cell invasion is MET mediated. We stimulated A549 cells with EGF alone or with erlotinib or the MET TKI, PHA-665752. EGF induced a twofold increase in cell invasion compared with control (P <0.05; Figure 4b). The addition of erlotinib or PHA-665752 reduced the number of invading cells to control levels, indicating that EGFR-driven cell invasion is MET dependent. In a similar experiment using HCC827 cells, in which EGFR is

constitutively activated, EGF stimulation did not increase tumor cell invasiveness compared with control levels; however, pharmacological inhibition of EGFR or MET activation significantly reduced the number of invading cells (Figure 4c; P <0.05).

To further elucidate the mechanism by which EGFR-activating mutations drive tumor cell invasion, we transfected HCC827 cells with EGFR-, HIF-1 α -, or MET-targeting siRNA and evaluated cell invasion. Knockdown of EGFR, HIF-1 α , or MET resulted in decreased invasive capacity (Figure 4d; P <0.001),

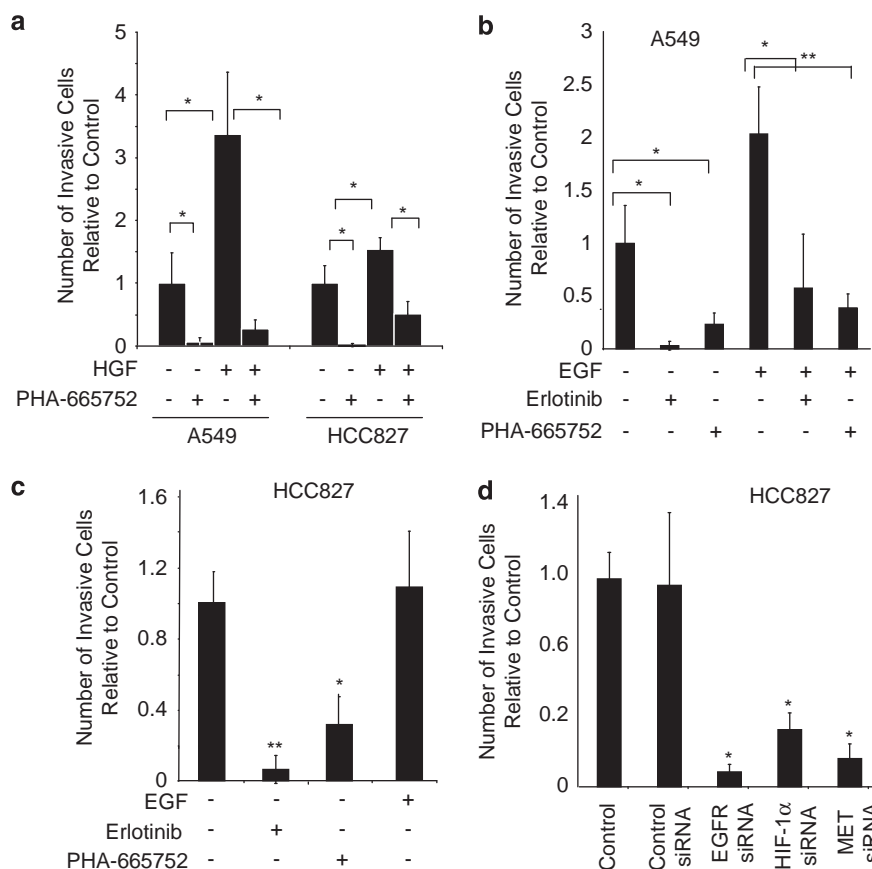


Figure 4 MET is required for EGFR-driven invasiveness of NSCLC cell lines. (a) A549 and HCC827 cells were seeded onto Matrigel-coated invasion chambers. Cells were treated with HGF (40 ng/ml) with or without PHA-663225 (1 μ M) and incubated for 48 h. Migrating cells were quantified by bright-field microscopy. Bars, % s.d.; * P <0.05. (b) A549 cells were seeded onto Matrigel-coated invasion chambers. Cells were treated with EGF (60 ng/ml), erlotinib (1 μ M), or PHA-663225 (1 μ M) and incubated for 24 h. EGF significantly enhanced cell invasion, and EGFR and MET inhibitors abrogated EGF-induced invasion. Bars, % s.d.; * P <0.05; ** P <0.001. (c) HCC827 cells were treated with EGF, erlotinib, or PHA-663225 and tumor cell invasion was evaluated by Boyden chamber assay. Bars, % s.d.; * P <0.05; ** P <0.001. (d) HCC827 cells were transfected with control siRNA or siRNA-targeting EGFR, HIF-1 α , or MET, and tumor cell invasion was evaluated by Boyden chamber assay. Bars, % s.d.; * P <0.001.

whereas control siRNA did not affect invasive capacity. To confirm that decreases in the number of invasive cells after erlotinib or siRNA treatment was indeed do to changes in the invasive capacity of tumor cells and not because of changes in cell viability or proliferation, we separately performed a similar study in which the number of invasive cells was normalized to the number of cells that did not invade through the chamber (Supplemental Figure 1). These data were in agreement with the findings shown above.

To determine whether MET inhibition impacted cellular growth, we conducted MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt) assays in the presence of increasing concentrations of erlotinib or PHA-665752. HCC827 cells were sensitive to the EGFR inhibitor erlotinib (IC_{50} <10 nM), whereas H1973 and A549 cells were less sensitive (Supplemental Figure 2). None of the three cell lines were sensitive to PHA-665752 at concentrations as high as 10 μ M. This suggested that invasiveness but not cell survival was MET dependent in these cell lines.

Hypoxia-induced invasiveness of NSCLC cells is MET dependent

Previous studies show that hypoxia increases the invasive capacity of tumor cells (Cuvier *et al.*, 1997). Therefore, we investigated whether MET activation mediates hypoxia-induced invasiveness of tumor cells. A549 cells were plated in Matrigel-coated Boyden chambers and incubated in normoxia (21% oxygen) or hypoxia (0.1% oxygen) with or without erlotinib or PHA-665752. Hypoxia enhanced tumor cell invasion more than 2.5-fold (P <0.001). Erlotinib caused a moderate decrease in tumor cell invasiveness, whereas MET inhibition reduced hypoxia-induced tumor cell invasion to levels below baseline (Figure 5a; P <0.001).

To examine the role of HIF-1 α in this pathway, we stably transfected A549 cells with a variant form of HIF-1 α (HA-HIF-1 α P402A;P564A) that is constitutively stabilized in normoxia because of proline to alanine substitutions at the VHL binding site critical for HIF-1 α polyubiquitination and degradation (Masson *et al.*, 2001; Hu *et al.*, 2003; Kim *et al.*, 2006). Expression of the HIF-1 α variant augmented MET production

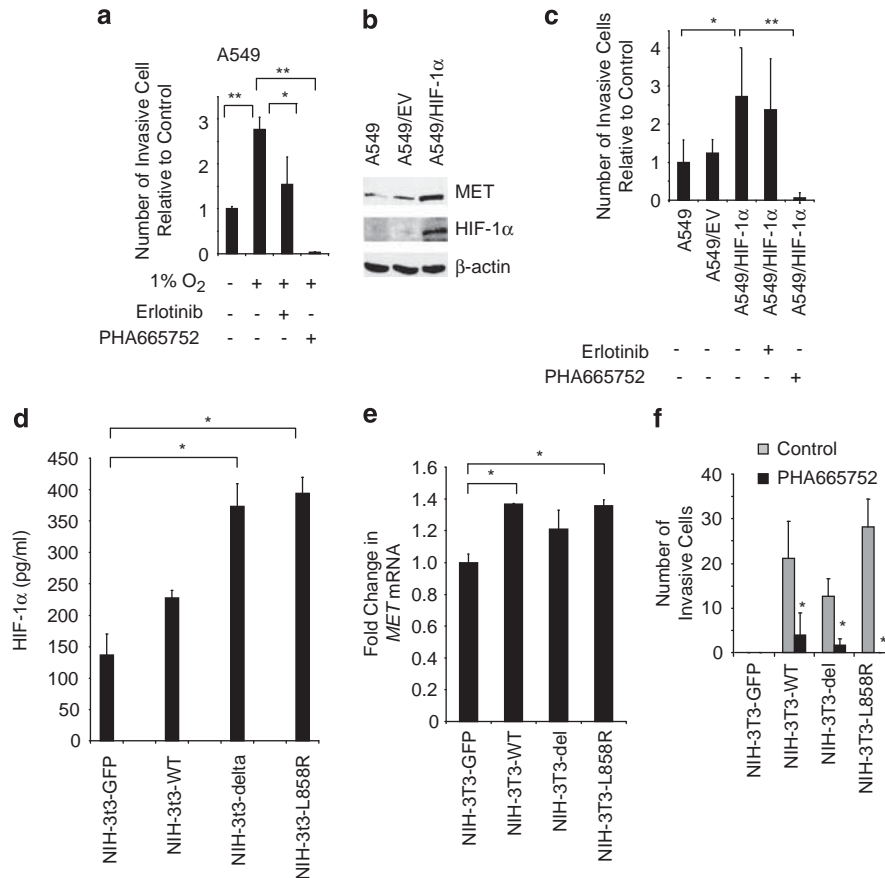


Figure 5 Hypoxia-induced tumor cell invasion is MET dependent. (a) A549 cells were seeded onto growth factor-reduced Matrigel-coated invasion chambers and incubated in normoxia or hypoxic (1% O₂) conditions with or without erlotinib (1 μ M) or PHA-663225 (1 μ M) for 24 h, and the number of invading cells was quantified. Hypoxia-induced invasion was MET dependent but EGFR independent. Bars, % s.d.; * P < 0.05; ** P < 0.001. (b) A549 cells were transfected to overexpress a degradation-resistant HIF-1 α variant (HA-HIF-1 α P402A;P564A), as shown by immunoblot. Increased HIF-1 α was associated with enhanced MET expression. (c) A549 cells and A549 cells transfected with empty vector (EV) or plasmid containing the HIF1 α mutant were evaluated for changes in invasive capacity in media with or without erlotinib (1 μ M) or PHA-663225 (1 μ M). After 24 h the number of invading cells was quantified. Bars, % s.d.; * P < 0.05; ** P = 0.006. (d) NIH-3T3 cells expressing GFP, WT EGFR, or EGFR bearing the L858R mutation or the deletion mutant Δ L747-S752del were evaluated for HIF-1 α levels by ELISA. * P < 0.05. (e) MET mRNA expression was evaluated by RT-PCR in NIH-3T3 cells expressing WT EGFR and EGFR-activating mutations. * P < 0.05. (f) NIH-3T3 cells expressing GFP, WT EGFR, or EGFR bearing the L858R mutation or the deletion mutant Δ L747-S752del were allowed to migrate through Matrigel-coated invasion chambers for 24 h with or without PHA-663225 (1 μ M). Bars, s.d.; * P < 0.05.

(Figure 5b) and was associated with enhanced tumor cell invasion (P < 0.05; Figure 5c). PHA-665752, but not erlotinib, reduced the number of invading cells to below baseline values (P < 0.05). These findings were consistent with MET being a downstream mediator of HIF-1 α -mediated invasiveness in these cells.

We examined the impact of WT and mutated EGFR, and the role of MET, on invasiveness in NIH-3T3 cells. NIH-3T3 cells were stably transfected with control GFP plasmid, WT *EGFR* or *EGFR* bearing the L858R mutation or the deletion mutant Δ L747-S752del. A greater than twofold increase in HIF-1 α levels was observed in cells bearing the Δ L747-S752del deletion (P = 0.02) or the L858R mutation (P = 0.01) compared with GFP transfected controls (Figure 5d). RT-PCR analysis revealed that NIH-3T3 cells expressing L858R expressed increased *MET* mRNA levels compared with GFP transfected controls (P = 0.02; Figure 5e). NIH-3T3 cells stably transfected with control GFP plasmid,

WT *EGFR*, or *EGFR* bearing the L858R mutation or the deletion mutant Δ L747-S752del were allowed to invade Matrigel-coated Boyden chambers with or without PHA-665752. Cells expressing WT or activated EGFR had enhanced invasive capacity compared with cells transfected with GFP vector, and inhibition of MET significantly reduced the number of invasive cells to near baseline values (Figure 5f).

Discussion

This study offers new insights into the mechanisms by which EGFR mediates its tumorigenic effects and provides new evidence that the HIF-1 α /MET axis is critical to regulating invasiveness induced not only by hypoxia but by EGFR as well, thus illustrating the convergence of two pathways known to drive invasive

tumor growth. In NSCLC cells, we showed that EGFR activation, by either EGFR kinase mutations or ligand binding, increased MET levels through a hypoxia-independent mechanism involving expression of HIF-1 α . MET was uncoupled from EGFR regulation, however, in a cell line with *MET* amplification, a finding consistent with the recently described role of *MET* amplification in EGFR TKI resistance (Engelman *et al.*, 2007). Overexpression of a constitutively active form of HIF-1 α also abrogated the regulation of MET levels by EGFR. Therefore, though this study shows that EGFR signaling can regulate MET levels and that MET can be downstream mediator of EGFR-induced invasiveness, it also suggests that there are ways by which this pathway may be bypassed.

We initially investigated MET levels in tumor specimens from 202 NSCLC patients by immunohistochemistry and observed increased levels of MET in tumors with *EGFR*-activating mutations compared with tumors with WT *EGFR*. Consistent with these findings, we observed elevated levels of MET in a previously described transgenic murine model of NSCLC with lung-specific expression of an *EGFR*-activating mutation. MET levels decreased significantly after treatment with the EGFR inhibitor erlotinib. We also observed that NSCLC cell lines expressing mutated *EGFR* exhibited elevated *MET* gene expression and protein levels compared with cells with WT *EGFR*, and these levels could be reduced by pharmacologically inhibiting EGFR or with siRNA directed against EGFR. The addition of an EGFR inhibitor decreased *MET* mRNA, indicating that in NSCLC cells, *EGFR*-activating mutations augment MET expression at the transcriptional level. Collectively, these results provide evidence that activated EGFR has a critical role in regulating MET expression in NSCLC tumor cells.

MET amplification has been described in the setting of gastric cancers (Smolen *et al.*, 2006) and NSCLC (Engelman *et al.*, 2007). Engelman *et al.* (2007) reported that among lung cancers with *EGFR*-activating mutations, *MET* amplification occurred in 22% of tumors that developed resistance to gefitinib or erlotinib. Overall, *MET* amplification occurs in only 4% of all NSCLC cases (Zhao *et al.*, 2005), whereas high levels of MET and p-MET are detectable in 36 and 21% of NSCLC cases, respectively (Nakamura *et al.*, 2007). MET expression also has been shown to be regulated by hypoxia (Pennacchietti *et al.*, 2003). Here, we found evidence that EGFR is a key regulator of MET levels in cells without *MET* amplification, and that this occurs through hypoxia-independent regulation of HIF-1 α . By contrast, in H1993 cells bearing *MET* gene amplification, EGFR blockade did not result in a reduction in *MET* levels, suggesting that *MET* amplification resulted in an uncoupling of MET protein levels from EGFR regulation.

Ligand-induced phosphorylation of EGFR has been shown to induce rises in HIF-1 α in a cell type-specific manner. We observed that HIF-1 α levels were elevated in NSCLC cell lines bearing *EGFR* mutations even in the absence of added ligand, and that treatment with an EGFR inhibitor diminished HIF-1 α expression.

Although this study did not specifically address the regulation of angiogenic factors such as VEGF, these findings are consistent with recent studies that EGFR regulates angiogenic factors, at least in part, through HIF-1 α -dependent mechanisms (Swinson *et al.*, 2004; Luwor *et al.*, 2005; Pore *et al.*, 2006; Swinson and O'Byrne, 2006). Studies designed to elucidate the mechanism by which EGFR-activating mutations regulate HIF-1 α levels are ongoing.

In addition to enhancing MET levels, EGFR activation resulted in increases in MET receptor phosphorylation within 30 min of EGF stimulation, an effect blocked by erlotinib. Similar observations have been made with other cell types (Bergstrom *et al.*, 2000; Jo *et al.*, 2000); these and other published data support the idea that EGFR may directly phosphorylate MET (Bergstrom *et al.*, 2000; Jo *et al.*, 2000). Hypoxia is a known regulator of HGF, presumably through HIF-1 α (Ide *et al.*, 2006). It is feasible that EGFR-activating mutations promote HGF production through HIF-1 α . Collectively, these data suggest that EGFR/HIF-1 α activation may not only regulate MET levels, but may also impact MET signal transduction through other mechanisms.

We investigated the consequences of EGFR-regulated MET. We observed that the invasiveness (Figure 4) but not survival (Supplemental Figure 2) of NSCLC cells bearing EGFR-activating mutations was MET dependent, as pharmacological inhibition or siRNA directed against MET abrogated cell invasion. Invasiveness and MET levels were reduced by siRNA knockdown of HIF-1 α , whereas EGFR levels were unaffected; indicating that MET is downstream of HIF-1 α and EGFR. We observed that EGF- and hypoxia-induced invasiveness were both MET dependent, and that heterologous expression of a constitutively active form of HIF-1 α induced invasiveness that was independent of EGF stimulation but remained MET dependent (Figure 5). Furthermore, heterologous expression of wild-type or mutated EGFR in NIH 3T3 fibroblasts increased invasiveness in a MET-dependent manner, providing further evidence that EGFR-mediated invasiveness is mediated at least in part by MET. These results support a model in which either hypoxia or EGFR activation can drive invasiveness by converging on a common HIF/MET pathway, which appears to be separable from EGFR-induced survival and proliferation. *MET* amplification appears to provide one route for circumventing this pathway. Others will likely emerge, and these findings do not exclude the likelihood that other pathways contribute to the invasive phenotype.

Our findings that EGFR can regulate MET levels through hypoxia-independent regulation of HIF-1 α , and that MET is a downstream mediator of both EGFR- and hypoxia-induced invasiveness, have important clinical and biological implications. Even for tumors thought to be primarily driven by the EGFR pathway (that is with activating EGFR mutations), targeting of the MET pathway in combination with EGFR blockade may further reduce tumor invasiveness

beyond the effect of EGFR inhibition alone, in addition to the previously noted potential benefit of preventing the emergence of resistance through MET amplification (Engelman *et al.*, 2007). It also raises the question of whether other tyrosine kinases may regulate MET in a similar manner in NSCLC and other diseases. Finally, the observation that the EGFR and hypoxia converge on the HIF-1 α /MET axis suggests that there may be additional overlap in the mechanisms by which EGFR and hypoxia promote malignant behavior and therapeutic resistance.

Materials and methods

Cell lines

Drs John Minna and Adi Gazdar (UT Southwestern Medical School, Dallas, TX, USA) provided H3255, H1975, H1993, and HCC827 cells. A549 and Calu-6 were obtained from the ATCC (Rockville, MD, USA). NIH-3T3 cells expressing WT *EGFR* or *EGFR* bearing the L858R mutation or the deletion mutant Δ L747-S752del (Shimamura *et al.*, 2005) were obtained from Dr Jeffrey Engelman (Dana-Farber Cancer Institute) and were maintained in 10% FBS DMEM containing 1 mg/ml puromycin.

Mice

Animals were treated in accordance with the guidelines of the US Department of Agriculture and the NIH. *Kras*^{LA1} mice (Johnson *et al.*, 2001) were obtained from Dr Tyler Jacks (Massachusetts Institute of Technology, Cambridge, MA, USA). *CCSP-rtTA* transgenic mice (obtained from J Whitsett at The University of Cincinnati, Cincinnati, OH, USA) were bred with *Tet-op-hEGFR L858R-Luc* to yield mice with lung tumors driven by EGFR activation (Ji *et al.*, 2006). At 6 months of age, lungs were collected. Animals bearing EGFR-driven lung tumors were treated with vehicle (1% Tween-80; Sigma-Aldrich, St Louis, MO, USA), or erlotinib (50 mg/kg/day) by gavage for 48 h.

Gene expression analysis

Affymetrix GeneChip Human Genome U133A (HG-U133A) was used to perform gene expression analysis on 53 gene arrays of NSCLC cell lines prepared by John Minna and colleagues (UT Southwestern, Dallas, TX, USA; Zhou *et al.*, 2006). CEL-type data files were obtained from NCBI-GEO dataset GSE4824 (NCBI-GEO, 2007). CHIP (2007) software (<http://biosun1.harvard.edu/complab/dchip/>) was used to generate probe-level gene expression, median intensity, percentage of probe set outliers, and percentage of single probe outliers (Lin *et al.*, 2004). Information files, including the HG-U133A gene information files and Chip Description Files, were downloaded from the Affymetrix web site. CEL and other data files were extracted. Array images were inspected for contamination and bad hybridization. Normalization was performed using the invariant-set normalization method (Li and Hung Wong, 2001). Model-based expression and background subtraction using the 5th percentile of region (perfect match only) was completed by checking for single, array, and probe outliers. In the array analysis and clustering, array outliers were treated as missing values and no log transformation was performed. Comparison within dCHIP of the WT *EGFR* vs mutated *EGFR* groups using a more than 1.2-fold change in gene expression, a 90% confidence interval for fold change, and a 90% present call

yielded one probeset for *MET* (203510_at). Data were further analyzed using GraphPad software (version 5, GraphPad Software Inc., La Jolla, CA, USA).

Detection of HIF-1 α , MET, and EGFR

Protein lysates were extracted using RIPA buffer (50mM Tris-HCl, pH 7.4, 150mM NaCl, 1mM EDTA, 1% Triton X-100, 1% sodium deoxycholate, and 0.1% SDS, and protease inhibitors. Protein (60 μ g) was used for western blotting. Antibodies against EGFR (Y1068, Cell Signaling Technology Inc., Danvers, MA, USA), EGFR (Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) MET (Santa Cruz), HIF-1 α (Pharmingen, San Diego, CA, USA), β -actin (Sigma), and vinculin (Sigma) were used. HIF-1 α , p-EGFR, MET, and p-MET ELISAs were obtained from R&D systems (Minneapolis, MN, USA).

RNA isolation and RT-PCR

HCC827 cells were treated with complete media with or without 1 μ M erlotinib for 12 h. Total RNA was extracted using Trizol (Life Technologies, Carlsbad, CA, USA), and purified with the RNeasy kit (Qiagen, Hilden, Germany). We used SuperScriptTM III RNase H-Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) to convert RNA into cDNA. The oligonucleotide primers used were published previously (Shimazaki *et al.*, 2003).

Plasmids and transfections

HIF-1 α cDNA (OriGene Technologies Inc., Rockville, MD, USA) was subcloned into the pcDNA3.1 vector with a flag tagged in the N-terminal, and the HIF-1 α mutant with proline to alanine substitutions positions 402 and 564 (HA-HIF-1 α P402A;P564A), which are known VHL-binding sites, was constructed as described (Kim *et al.*, 2006). This form is stabilized in normoxia because of the loss of VHL-mediated polyubiquitination and subsequent degradation (Masson *et al.*, 2001; Hu *et al.*, 2003). For siRNA transfections, HCC827 cells were transfected with siRNA targeting EGFR, MET, HIF-1 α and control siRNA at a final concentration of 100 nM using Dharmafect 1 transfection reagent (Dharmacon, Lafayette, CO, USA). Protein was isolated after 72 h. Transfected cells were plated for invasion assays after 48 h.

Invasion assay

We seeded 2.5×10^4 cells in the upper chamber of 24-well BD Biocoat growth factor reduced Matrigel invasion chambers (8.0 μ m pore, Becton Dickinson, Bedford, MA, USA) with 0% FBS media and added media containing 10% FBS to the lower chamber. After 24 h, cells in the upper chamber were removed by scraping. Cells that migrated to the lower chamber were stained and counted using bright-field microscopy under a low-power ($\times 40$) objective. PHA-663225 was obtained from Pfizer (New York, NY, USA).

Clinical specimens

Tissue specimens from 202 surgically resected lung carcinomas were obtained from the Lung Cancer Specialized Program of Research Excellence (SPORE) Tissue Bank at The University of Texas MD Anderson Cancer Center (Houston, TX, USA). Two hundred and two specimens had known EGFR status. Microarrays for each specimen were created with three cores from formalin-fixed, paraffin-embedded blocks. All specimens were of pathologic TNM stages I–IV according to the revised International System for Staging Lung Cancer (Beaumont and Srean, 2003).

Immunohistochemistry staining

Using paraffin-embedded tissue sections, antigen retrieval was performed by steaming in citrate buffer (pH 6.0). Endogenous peroxidases were blocked using 3% H₂O₂. After protein blocking, slides were incubated with anti-HGFR (1:50; R&D Systems), washed, and incubated with a Universal LSAB + Kit/HRP, visualization kit (DakoCytomation, Carpinteria, CA, USA). For tumor sections from transgenic animals, antigen retrieval and blocking was performed as above. Slides were incubated in 1:100 anti-mouse MET antibodies (Santa Cruz) and then in secondary antibody (Jackson Research Laboratories, Bar Harbor, ME, USA). NSCLC specimens were used as positive controls for MET staining. As a negative control, we followed the above procedure omitting the primary antibodies. For quantification, each specimen was evaluated using an intensity score (0, 1, 2, or 3) and an extension percentage (Yang *et al.*, 2008). The final staining score was the product of these two values. An average from the three cores was obtained for each specimen.

Statistics

Student's *t*-tests were performed using two-tailed tests with unequal variance for Gaussian distributed data. For statistical analysis of clinical specimens, Wilcoxon rank-sum tests were used when comparing continuous variables between mutation groups. To correlate mutation and other discrete covariates,

we used a chi-square test or Fisher's exact test. Two-sided *P*-values ≤ 0.05 were considered significant.

Accession numbers

SNP and CGH raw data are available in the Gene Expression Omnibus (GEO) database: GEO accession GSE4824 (<http://www.ncbi.nlm.nih.gov/geo/>).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Drs Mien-Chie Hung, Scott Lippman, and Kian Ang for their critical review, Dr Jeffrey Engelman for the EGFR-transfected NIH 3T3 cell lines, and Joseph Munch for editorial assistance. Supported in part by the NIH Lung Cancer SPORE grant P50 CA70907, NIH Grant P01 CA06294, Department of Defense Grant W81XWH-07-1-0306 01, awards from the Metastasis Foundation, the Physician-Scientist Program, and the American Society for Clinical Oncology Career Development Award (JVH). JVH is a Damon Runyon-Lilly Clinical Investigator supported in part by the Damon Runyon Cancer Research Foundation (CI 24-04).

References

- Beadsmoore CJ, Screaton NJ. (2003). Classification, staging and prognosis of lung cancer. *Eur J Radiol* **45**: 8–17.
- Bergstrom JD, Westermarck B, Heldin NE. (2000). Epidermal growth factor receptor signaling activates met in human anaplastic thyroid carcinoma cells. *Exp Cell Res* **259**: 293–299.
- Bredin CG, Liu Z, Klominek J. (2003). Growth factor-enhanced expression and activity of matrix metalloproteases in human non-small cell lung cancer cell lines. *Anticancer Res* **23**: 4877–4884.
- Ciardiello F, Tortora G. (2008). EGFR antagonists in cancer treatment. *N Engl J Med* **358**: 1160–1174.
- Cooper CS, Tempest PR, Beckman MP, Heldin CH, Brookes P. (1986). Amplification and overexpression of the met gene in spontaneously transformed NIH3T3 mouse fibroblasts. *EMBO J* **5**: 2623–2628.
- Cuvier C, Jang A, Hill RP. (1997). Exposure to hypoxia, glucose starvation and acidosis: effect on invasive capacity of murine tumor cells and correlation with cathepsin (L + B) secretion. *Clin Exp Metastasis* **15**: 19–25.
- Damstrup L, Rude Voldborg B, Spang-Thomsen M, Brunner N, Skovgaard Poulsen H. (1998). *In vitro* invasion of small-cell lung cancer cell lines correlates with expression of epidermal growth factor receptor. *Br J Cancer* **78**: 631–640.
- Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO *et al.* (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**: 1039–1043.
- Forsythe JA, Jiang BH, Iyer NV, Agani F, Leung SW, Koos RD *et al.* (1996). Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Mol Cell Biol* **16**: 4604–4613.
- Fukuoka M, Yano S, Giaccone G, Tamura T, Nakagawa K, Douillard JY *et al.* (2003). Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial) [corrected]. *J Clin Oncol* **21**: 2237–2246.
- Guo A, Villen J, Kornhauser J, Lee KA, Stokes MP, Rikova K *et al.* (2008). Signaling networks assembled by oncogenic EGFR and c-Met. *Proc Natl Acad Sci USA* **105**: 692–697.
- Hamada J, Nagayasu H, Takayama M, Kawano T, Hosokawa M, Takeichi N. (1995). Enhanced effect of epidermal growth factor on pulmonary metastasis and *in vitro* invasion of rat mammary carcinoma cells. *Cancer Lett* **89**: 161–167.
- Hirami Y, Aoe M, Tsukuda K, Hara F, Otani Y, Koshimune R *et al.* (2004). Relation of epidermal growth factor receptor, phosphorylated-Akt, and hypoxia-inducible factor-1 α in non-small cell lung cancers. *Cancer Lett* **214**: 157–164.
- Hu CJ, Wang LY, Chodosh LA, Keith B, Simon MC. (2003). Differential roles of hypoxia-inducible factor 1 α (HIF-1 α) and HIF-2 α in hypoxic gene regulation. *Mol Cell Biol* **23**: 9361–9374.
- Huang PH, Mukasa A, Bonavia R, Flynn RA, Brewer ZE, Cavenee WK *et al.* (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proc Natl Acad Sci USA* **104**: 12867–12872.
- Ichimura E, Maeshima A, Nakajima T, Nakamura T. (1996). Expression of c-met/HGF receptor in human non-small cell lung carcinomas *in vitro* and *in vivo* and its prognostic significance. *Jpn J Cancer Res* **87**: 1063–1069.
- Ide T, Kitajima Y, Miyoshi A, Ohtsuka T, Mitsuno M, Ohtaka K *et al.* (2006). Tumor-stromal cell interaction under hypoxia increases the invasiveness of pancreatic cancer cells through the hepatocyte growth factor/c-Met pathway. *Int J Cancer* **119**: 2750–2759.
- Janne PA, Engelman JA, Johnson BE. (2005). Epidermal growth factor receptor mutations in non-small-cell lung cancer: implications for treatment and tumor biology. *J Clin Oncol* **23**: 3227–3234.
- Ji H, Li D, Chen L, Shimamura T, Kobayashi S, McNamara K *et al.* (2006). The impact of human EGFR kinase domain mutations on lung tumorigenesis and *in vivo* sensitivity to EGFR-targeted therapies. *Cancer Cell* **9**: 485–495.

- Jo M, Stolz DB, Espen JE, Dorko K, Michalopoulos GK, Strom SC. (2000). Cross-talk between epidermal growth factor receptor and c-Met signal pathways in transformed cells. *J Biol Chem* **275**: 8806–8811.
- Johnson L, Mercer K, Greenbaum D, Bronson RT, Crowley D, Tuveson DA et al. (2001). Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**: 1111–1116.
- Kim WY, Safran M, Buckley MR, Ebert BL, Glickman J, Bosenberg M et al. (2006). Failure to prolyl hydroxylate hypoxia-inducible factor α phenocopies VHL inactivation *in vivo*. *EMBO J* **25**: 4650–4662.
- Kris MG, Natale RB, Herbst RS, Lynch Jr TJ, Prager D, Belani CP et al. (2003). Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. *JAMA* **290**: 2149–2158.
- Li C, Hung Wong W. (2001). Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol* **2**: RESEARCH0032.
- Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, Li C. (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**: 1233–1240.
- Liu C, Tsao MS. (1993). *In vitro* and *in vivo* expressions of transforming growth factor- α and tyrosine kinase receptors in human non-small-cell lung carcinomas. *Am J Pathol* **142**: 1155–1162.
- Lutterbach B, Zeng Q, Davis LJ, Hatch H, Hang G, Kohl NE et al. (2007). Lung cancer cell lines harboring MET gene amplification are dependent on Met for growth and survival. *Cancer Res* **67**: 2081–2088.
- Luwor RB, Lu Y, Li X, Mendelsohn J, Fan Z. (2005). The antiepidermal growth factor receptor monoclonal antibody cetuximab/C225 reduces hypoxia-inducible factor-1 α , leading to transcriptional inhibition of vascular endothelial growth factor expression. *Oncogene* **24**: 4433–4441.
- Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW et al. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* **350**: 2129–2139.
- Masson N, Willam C, Maxwell PH, Pugh CW, Ratcliffe PJ. (2001). Independent function of two destruction domains in hypoxia-inducible factor- α chains activated by prolyl hydroxylation. *EMBO J* **20**: 5197–5206.
- Nakamura Y, Niki T, Goto A, Morikawa T, Miyazawa K, Nakajima J et al. (2007). c-Met activation in lung adenocarcinoma tissues: An immunohistochemical analysis. *Cancer Sci* **98**: 1006–1013.
- NCBI-GEO (2007). <http://www.ncbi.nlm.nih.gov/geo/>.
- Olivero M, Rizzo M, Madeddu R, Casadio C, Pennacchietti S, Nicotra MR et al. (1996). Overexpression and activation of hepatocyte growth factor/scatter factor in human non-small-cell lung carcinomas. *Br J Cancer* **74**: 1862–1868.
- Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**: 1497–1500.
- Pai R, Nakamura T, Moon WS, Tarnawski AS. (2003). Prostaglandins promote colon cancer cell invasion; signaling by cross-talk between two distinct growth factor receptors. *FASEB J* **17**: 1640–1647.
- Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I et al. (2004). EGF receptor gene mutations are common in lung cancers from 'never smokers' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA* **101**: 13306–13311.
- Pao W, Miller VA. (2005). Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. *J Clin Oncol* **23**: 2556–2568.
- Park M, Dean M, Kaul K, Braun MJ, Gonda MA, Vande Woude G. (1987). Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors. *Proc Natl Acad Sci USA* **84**: 6379–6383.
- Peng XH, Karna P, Cao Z, Jiang BH, Zhou M, Yang L. (2006). Cross-talk between epidermal growth factor receptor and hypoxia-inducible factor-1 α signal pathways increases resistance to apoptosis by up-regulating survivin gene expression. *J Biol Chem* **281**: 25903–25914.
- Pennacchietti S, Michieli P, Galluzzo M, Mazzone M, Giordano S, Comoglio PM. (2003). Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer Cell* **3**: 347–361.
- Peruzzi B, Bottaro DP. (2006). Targeting the c-Met signaling pathway in cancer. *Clin Cancer Res* **12**: 3657–3660.
- Phillips RJ, Mestas J, Gharraee-Kermani M, Burdick MD, Sica A, Belperio JA et al. (2005). Epidermal growth factor and hypoxia-induced expression of CXC chemokine receptor 4 on non-small cell lung cancer cells is regulated by the phosphatidylinositol 3-kinase/PTEN/AKT/mammalian target of rapamycin signaling pathway and activation of hypoxia inducible factor-1 α . *J Biol Chem* **280**: 22473–22481.
- Pore N, Jiang Z, Gupta A, Cerniglia G, Kao GD, Maity A. (2006). EGFR tyrosine kinase inhibitors decrease VEGF expression by both hypoxia-inducible factor (HIF)-1-independent and HIF-1-dependent mechanisms. *Cancer Res* **66**: 3197–3204.
- Qiao H, Hung W, Tremblay E, Wojcik J, Gui J, Ho J et al. (2002). Constitutive activation of met kinase in non-small-cell lung carcinomas correlates with anchorage-independent cell survival. *J Cell Biochem* **86**: 665–677.
- Rosario M, Birchmeier W. (2003). How to make tubes: signaling by the Met receptor tyrosine kinase. *Trends Cell Biol* **13**: 328–335.
- Scheving LA, Stevenson MC, Taylormoore JM, Traxler P, Russell WE. (2002). Integral role of the EGF receptor in HGF-mediated hepatocyte proliferation. *Biochem Biophys Res Commun* **290**: 197–203.
- Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S et al. (2005). Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* **353**: 123–132.
- Shimamura T, Lowell AM, Engelman JA, Shapiro GI. (2005). Epidermal growth factor receptors harboring kinase domain mutations associate with the heat shock protein 90 chaperone and are destabilized following exposure to geldanamycins. *Cancer Res* **65**: 6401–6408.
- Shimazaki K, Yoshida K, Hirose Y, Ishimori H, Katayama M, Kawase T. (2003). Cytokines regulate c-Met expression in cultured astrocytes. *Brain Res* **962**: 105–110.
- Siegfried JM, Weissfeld LA, Luketich JD, Weyant RJ, Gubish CT, Landreneau RJ. (1998). The clinical significance of hepatocyte growth factor for non-small cell lung cancer. *Ann Thorac Surg* **66**: 1915–1918.
- Smolen GA, Sordella R, Muir B, Mohapatra G, Barmettler A, Archibald H et al. (2006). Amplification of MET may identify a subset of cancers with extreme sensitivity to the selective tyrosine kinase inhibitor PHA-665752. *Proc Natl Acad Sci USA* **103**: 2316–2321.
- Stabile LP, Lyker JS, Huang L, Siegfried JM. (2004). Inhibition of human non-small cell lung tumors by a c-Met antisense/U6 expression plasmid strategy. *Gene Ther* **11**: 325–335.
- Swinson DE, Jones JL, Cox G, Richardson D, Harris AL, O'Byrne KJ. (2004). Hypoxia-inducible factor-1 α in non small cell lung cancer: relation to growth factor, protease and apoptosis pathways. *Int J Cancer* **111**: 43–50.
- Swinson DE, O'Byrne KJ. (2006). Interactions between hypoxia and epidermal growth factor receptor in non-small-cell lung cancer. *Clin Lung Cancer* **7**: 250–256.
- Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, von Pawel J et al. (2005). Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *Lancet* **366**: 1527–1537.
- Weinberger PM, Yu Z, Kowalski D, Joe J, Manger P, Psyrri A et al. (2005). Differential expression of epidermal growth factor receptor, c-Met, and HER2/neu in chordoma compared with 17 other malignancies. *Arch Otolaryngol Head Neck Surg* **131**: 707–711.

- Yang Y, Wislez M, Fujimoto N, Prudkin L, Izzo JG, Uno F *et al.* (2008). A selective small molecule inhibitor of c-Met, PHA-665752, reverses lung premalignancy induced by mutant K-ras. *Mol Cancer Ther* 7: 952–960.
- Yi S, Chen JR, Viallet J, Schwall RH, Nakamura T, Tsao MS. (1998). Paracrine effects of hepatocyte growth factor/scatter factor on non-small-cell lung carcinoma cell lines. *Br J Cancer* 77: 2162–2170.
- Zhao X, Weir BA, LaFramboise T, Lin M, Beroukhim R, Garraway L *et al.* (2005). Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. *Cancer Res* 65: 5561–5570.
- Zhong H, Chiles K, Feldser D, Laughner E, Hanrahan C, Georgescu MM *et al.* (2000). Modulation of hypoxia-inducible factor 1 α expression by the epidermal growth factor/phosphatidylinositol 3-kinase/PTEN/AKT/FRAP pathway in human prostate cancer cells: implications for tumor angiogenesis and therapeutics. *Cancer Res* 60: 1541–1545.
- Zhou BB, Peyton M, He B, Liu C, Girard L, Caudler E *et al.* (2006). Targeting ADAM-mediated ligand cleavage to inhibit HER3 and EGFR pathways in non-small cell lung cancer. *Cancer Cell* 10: 39–50.

Supplementary Information accompanies the paper on the Oncogene website (<http://www.nature.com/onc>)

**Stromal EGFR contributes to VEGF inhibitor resistance in murine models
of human NSCLC**

Tina Cascone^{1,4}, Matthew H. Herynk¹, Li Xu¹, Zhiqiang Du¹, Humam Kadara¹,
Monique B. Nilsson¹, Carol J. Oborn², Yun-Yong Park³, Ju-Seog Lee³, Fortunato
Ciardiello⁴, Robert R. Langley² and John V. Heymach^{1,2}.

¹Department of Thoracic/Head and Neck Medical Oncology; ²Cancer Biology;
³Systems Biology; The University of Texas M. D. Anderson Cancer Center,
Houston, TX, 77030 USA;

⁴Division of Medical Oncology, 'F. Magrassi - A. Lanzara' Department of Clinical
and Experimental Medicine, Second University of Naples, Naples, Italy

Corresponding Author: John V. Heymach

Departments of Thoracic/Head and Neck Medical
Oncology and Cancer Biology

The University of Texas M. D. Anderson Cancer
Center, Unit 432

1515 Holcombe Blvd, 77030 Houston, TX

Phone: 713-792-6363

Fax: 713-792-1220

Email: jheykach@mdanderson.org

Abstract

The mechanisms by which tumors develop resistance to angiogenesis inhibitors, and the relative contributions of tumor cells and stroma to resistance, have not been established. We developed human NSCLC xenograft models of resistance to the VEGF inhibitor bevacizumab in mice and, using species-specific profiling, investigated tumor cell and stromal mechanisms of resistance. Mouse- and human-specific profiling demonstrated that gene expression changes associated with acquired resistance occurred predominantly in stromal (mouse) and not tumor (human) cells. Components of the EGFR and FGFR2 pathways were significantly upregulated in stroma, but not in tumor cells. Increased phosphoEGFR was detected on pericytes of xenografts that acquired resistance and on endothelium of tumors with primary resistance. Acquired resistance was associated with a pattern of pericyte-covered, normalized revascularization, whereas tortuous, uncovered vessels were observed in primary resistance. Dual targeting of VEGF and EGFR pathways with bevacizumab and erlotinib, or the VEGFR/EGFR inhibitor vandetanib, reduced pericyte coverage and delayed resistance. These findings demonstrate that alterations in tumor stromal pathways, including EGFR and FGFR2, are associated with, and may contribute to VEGF inhibitor resistance and that targeting these pathways may improve efficacy. Understanding stromal signaling may be critical for developing biomarkers for angiogenesis inhibitors and improving combination regimens.

INTRODUCTION

Tumor growth and metastatic spread are dependent on the acquisition of a vascular supply, which occurs at least in part through angiogenesis (1-3). This process is regulated by the balance of pro- and anti-angiogenic factors (4). Therapeutic approaches to targeting tumor vasculature have focused largely on vascular endothelial growth factor (VEGF), the prototypical pro-angiogenic factor that stimulates endothelial cells to proliferate, migrate, and produce proteases necessary for the formation of new vasculature networks (5-9). Strategies to inhibit VEGF-induced angiogenesis have been developed, including anti-VEGF monoclonal antibodies, such as bevacizumab (Avastin; Genentech, South San Francisco, CA) (10, 11), and VEGF receptor (VEGFR) tyrosine kinase inhibitors (TKIs) (12).

Phase III trials have shown that the addition of bevacizumab (BV) to standard therapy prolonged progression-free or overall survival, and improved objective tumor responses, in patients with advanced malignancies, including non-small cell lung cancer (NSCLC) (13). Despite this initial benefit however, therapeutic resistance inevitably emerges in most patients. In patients with colorectal cancer, we observed that acquired resistance was associated with a rise in circulating proangiogenic factors in plasma including basic fibroblast growth factor (bFGF), hepatocyte growth factor (HGF) placental growth factor (PIGF), stromal-derived factor-1, (SDF1) as well as factors associated with myeloid recruitment (e.g. macrophage chemoattractant protein-3 (14). Preclinical studies of therapeutic resistance to angiogenesis inhibitors have described compensatory mechanisms

acquired from tumors in the face of anti-angiogenic therapy, including a hypoxia-induced upregulation of pro-angiogenic molecules, such as FGF family members, in tumors with acquired resistance to VEGFR pathway inhibitors (15). Genetic alterations in tumor cells, such as p53 loss, may also promote resistance to hypoxia-induced apoptosis, resulting in a decreased vascular dependence (16). Incomplete target inhibition following treatment with VEGFR antagonists has been described in orthotopic models of pancreatic cancer as well as in patients (17, 18). Moreover, while showing potent antitumor activity in mouse models of pancreatic neuroendocrine carcinoma and glioblastoma, VEGF pathway inhibitors concomitantly may also increase invasiveness and metastatic potential (19).

Collectively, these findings emphasize the adaptive nature of tumor cells to anti-angiogenic therapies, whereas the contribution of stromal cells remains unclear. In this study, we specifically investigated stromal changes associated with resistance to anti-VEGF therapy in two human NSCLC xenograft models of acquired and primary resistance to BV. We performed mouse- (stroma) and human (tumor cell)-specific gene expression profiling to further validate the role of specific stromal molecules involved in resistance. Our results demonstrate that acquired or primary resistance to VEGF inhibition is associated with distinct patterns of vascularization in NSCLC xenografts. The vast majority of gene expression changes observed in resistant tumors occurred in tumor stroma and not in tumor cells themselves, suggesting stromal changes are likely to play a role in the resistant phenotype. Among factors known to regulate angiogenesis,

overexpression of stromal EGFR and FGFR2 signaling pathways was observed in resistant tumors and blockade of the EGFR pathway reversed the increased pericyte coverage observed in vessels of resistant tumors. The results suggest that stromal mechanisms, including an upregulation of EGFR pathway, can contribute to the resistant phenotype, and that targeting these pathways may delay tumor growth and the onset of resistance.

RESULTS

H1975 and A549 Xenografts Exhibit Acquired and Primary Resistance to Anti-angiogenic Therapy

To investigate the mechanisms by which NSCLC xenografts develop resistance to VEGF blockade, we injected male nude mice with either H1975 or A549 human NSCLC adenocarcinoma cells. These models were selected because in prior studies (20) we observed that A549 xenograft tumors were relatively insensitive to VEGF signaling inhibitors including BV (primary resistance), whereas H1975 initially experienced significant tumor shrinkage typically lasting for more than one month. Furthermore, the tumor cells contain two common alterations associated with EGFR TKI resistance: a T790M EGFR mutation (H1975 model) (21) and KRas mutation (A549) (22). Approximately 3 weeks after tumor cell injection, mice bearing tumors with a mean volume of $\sim 270 \text{ mm}^3$ were randomized to receive either vehicle (control) or BV, as described in detail in the Materials and Methods section. Animals were treated for 2 weeks (short-term treatment) or until mice were sacrificed due to tumor burden. Tumors were

considered to be resistant when they tripled in volume (progression) compared with the pre-treatment tumor size and progression-free survival (PFS) was measured as the time from initiation of treatment until tumor progression. In H1975 tumors, 2 weeks of BV inhibited tumor growth by 77% compared to controls (Figure 1A) whereas in A549 xenografts the drug produced only a 16% reduction in tumor growth compared with vehicle-treated tumors (Figure 1B).

The individual tumor growth curves shown in Figure 1C&D illustrate the growth kinetics of H1975 and A549 xenografts treated with vehicle and BV for a longer term until progression. All H1975 vehicle-treated xenografts progressed within 31 days of the onset of treatment, showing a median PFS of 7 days. In contrast, 67% of xenografts receiving BV developed resistance, and the median PFS was 138 days, ($p < 0.001$, log-rank test) (Figure 1C). A549 tumors were less responsive to BV, showing a median PFS of 35 days compared with 22 days of vehicle-treated tumors ($p = 0.16$), (Figure 1D). Collectively, these results show that H1975 tumors are initially responsive to BV therapy but acquire resistance after prolonged treatment with the drug, whereas A549 tumors demonstrate primary resistance.

Acquired Resistance to BV Is Associated With Sustained Inhibition of VEGFR2 Activation

To determine whether acquired resistance to BV was due to increased VEGFR2 signaling, potentially through increased expression of murine VEGF or other mechanism to bypass blockade of human VEGF by this agent, we evaluated the

phosphorylation status of VEGFR2 in vehicle-treated and BV-resistant tumors using immunofluorescence (IF) staining. In control H1975 tumors, p-VEGFR2 was readily detected on CD31⁺ tumor-associated endothelial cells (ECs), but it was inhibited on vessels of BV-sensitive tumors (short-term treatment) and this result was sustained when tumors progressed while receiving BV therapy (Figure 2A). To evaluate changes in stromal (defined here as non-tumor cells derived from the host) and tumor-derived VEGF in H1975 BV-resistant tumors, we quantified mouse and human VEGFA mRNA expression by quantitative real-time PCR (qRT-PCR). We observed no change in mouse *Vegfa* mRNA expression in resistant xenografts compared with controls, whereas human *VEGFA* mRNA levels were increased in resistant tumors compared with controls ($p < 0.05$, Figure 2B), despite this ligand increase, VEGFR2 phosphorylation remained suppressed during BV treatment in resistant tumors.

We assessed whether the tumor growth inhibition observed in H1975 xenografts after short-term treatment with BV was associated with increased endothelial cell apoptosis. We performed double IF staining for CD31 (red) and TUNEL positivity (green) to identify apoptotic cells, (Figure 2C). The percentage of apoptotic ECs was significantly increased following 2 weeks of BV treatment compared to control xenografts ($p < 0.01$). However, at time of progression, H1975 xenografts showed a significant decrease in endothelial cell apoptosis compared with short-term treatment ($p < 0.05$) to levels comparable to control tumors (Figure 2D). Thus, endothelial cell apoptosis increased while tumors were initially responding to VEGF signaling blockade and returned to levels comparable with controls in

tumors that acquired BV-resistance. Furthermore, the percentage of total apoptotic (TUNEL⁺) cells showed a trend toward increase in H1975 BV-sensitive xenografts (BV 2 weeks) compared with controls (p=0.06), but not in tumors that progressed compared with controls (see Supplemental Figure 1A - B).

Stromal and Tumor Cell Gene Expression on Changes in H1975 BV-Resistant Xenografts

To identify changes in stromal and tumor gene expression associated with acquired resistance to anti-VEGF therapy, we performed RNA microarray analyses comparing H1975 control and BV-resistant xenografts (N=3 samples each group) using Illumina mouse (WG-6 v2) and human (WG-6 v3)-specific expression arrays. Probes in these arrays have been shown to minimize cross-species reactivity (Eun Sung Park, unpublished data). We found that a larger number of stromal mouse genes (1385) were significantly modulated in BV-resistant (BV progression) vs. control xenografts (vehicle progression) compared to human tumor genes (98), according to the statistic criteria described in Materials and Methods. We observed significant changes in the expression of genes involved in angiogenesis, lymphangiogenesis and hypoxia signaling pathway between BV-resistant and control xenografts. Both *Egfr* and *Fgfr2* genes were upregulated in the stromal compartment, but not in tumor cells, of H1975 BV-resistant tumors compared with controls, as well as stromal molecules and ligands associated with these signaling pathways (e.g., *Ep gn*, *Areg*, *Fgf13*, *Fgfbp1*) (Figure 3A, see Supplemental Table 1). Among human angiogenic

genes, CA9 (hypoxia-regulated gene) was significantly upregulated in BV-resistant tumors (Figure 3A, see Supplemental Table 2).

We next sought to identify genes predicted to be important in the phenotype of acquired resistance to the anti-angiogenic effect of BV through both their modulation in expression and molecular interactions. Functional gene-interaction network analyses of gene features differentially expressed between the mouse stroma of BV-resistant and vehicle-treated H1975 xenografts using Ingenuity Pathways Analysis (IPA) revealed the significant modulation in the predicted function of a gene neighborhood and interaction network surrounding *Egfr*, based on the number of focus genes and nodes of interaction ($p < 0.001$; Figure 3B). In addition, the modulated gene network associated with *Egfr* gene expression included down-regulated pro-apoptotic genes, such as the BH3-only family member protein, *Bax* and the apoptotic peptidase activating factor 1 (*Apaf1*). Genes with pro-survival functions were up-regulated, such as the heat shock protein *Dnajb1*.

Next, to validate the changes in expression of the significantly modulated network-hub gene, *Egfr*, we assessed the human and mouse mRNA levels using qRT-PCR. We observed a 2.5 fold increase in mouse *Egfr* mRNA levels in H1975 resistant xenografts compared with controls ($p < 0.05$), whereas human EGFR mRNA levels were not significantly different than controls (Figure 3C). We also validated the stromal expression of *Fgfr2*, which we noted to be upregulated in BV-resistant H1975 tumors in the microarray analysis. A significant increase in

mouse *Fgfr2* mRNA expression, but not human FGFR2, was observed in H1975 resistant xenografts compared with controls ($p < 0.05$, Figure 3D).

EGFR Is Activated on Stromal Cells of H1975 and A549 BV-Resistant Tumors

Given our observation that *Egfr* mRNA is increased in BV-resistant tumors, we next evaluated EGFR protein expression in H1975 tumors by IF staining using antibodies directed against CD31 (red) and total EGFR (green) (see Supplemental Figure 2A). We observed an increase in EGFR⁺ staining in H1975 BV-resistant tumors (BV progression) compared with controls (vehicle progression). Quantification of EGFR staining by laser scanning cytometry (LSC) analysis indicated a nearly 10-fold increase in the percentage of total EGFR-expressing cells in H1975 BV-resistant compared with control tumors ($p < 0.01$, Figure 4A). We also evaluated EGF protein expression changes by immunohistochemistry (IHC) in H1975 vehicle-and BV-treated xenografts at progression and noticed increased expression of the EGFR ligand in resistant tumors vs. controls (see Supplemental Figure 2B).

We next evaluated the EGFR activation status in H1975 and A549 xenografts following treatments with vehicle and BV at time of progression. Confocal microscopy was performed to analyze specimens stained with antibodies directed to CD31 (red) and p-EGFR (green). As shown in Figure 4B, BV resistance was associated with a marked increase in p-EGFR in both H1975 and A549 tumors compared with controls; however, notable differences in the

staining pattern were observed between the two xenograft models. In H1975, p-EGFR was increased on the vascular supporting cells (VSCs) of resistant tumors compared with controls ($p < 0.01$, Figure 4C, left), whereas, p-EGFR expression was increased on CD31⁺ ECs in A549 BV-resistant xenografts compared with controls ($p < 0.05$, Figure 4C, right).

To identify the population of VSCs expressing p-EGFR in H1975 BV-resistant tumors, we performed IF staining for p-EGFR (red) and desmin (green), a marker for pericytes (Figure 4D). We found p-EGFR expression to be significantly increased on desmin⁺ cells (pericytes) of BV-resistant H1975 xenografts compared with controls ($p < 0.01$, Figure 4E). Taken together our results suggest that increased stromal EGFR is associated with BV resistance, and that multiple stromal cell types can express EGFR.

bFGF/FGFR2 Overexpression in H1975 Xenografts Resistant to BV Therapy

Based on our observation that mouse *Fgfr2* gene expression was increased in the stromal compartment of BV-resistant H1975 tumors, we performed immunofluorescent co-localization studies on H1975 vehicle and BV-treated xenografts at progression ($n=4$, each group) using CD31 (red) and FGFR2 (green) antibodies (Figure 5A). We observed a significant increase in total FGFR2 expression in resistant tumors compared with controls ($p < 0.001$; Figure 5B). Furthermore, to assess changes in the FGFR2 ligand, we next measured the plasma concentration (pg/mL) and the expression levels of basic FGF (bFGF). We found a 1.5 fold increase in the levels of the circulating cytokine in

BV-resistant tumors compared with controls ($p=0.058$; Figure 5C). The immunohistochemical analysis of H1975 vehicle and BV-treated xenografts at progression ($n=4$, each group) suggested that BV-resistance is associated with increased expression of bFGF compared with controls (Figure 5D).

Resistance to BV is Associated with Tumor Revascularization and Morphological Changes in the Vasculature

Because the primary mechanism of action of BV is directed against blood vessels, we quantified the microvascular density (MVD) of H1975 and A549 xenografts. We noted a 3-fold MVD reduction in H1975 tumors treated with BV for 2 weeks compared with controls ($p<0.01$, see Supplemental Figure 3 - Figure 6B, left). Vessel density of A549 tumors treated for 2 weeks showed a trend towards a decrease compared with controls (see Supplemental Figure 3 - Figure 6B, right). To determine whether the vascular effects observed after 2 weeks of BV therapy persisted in tumors receiving BV long-term treatment, we quantified the MVD in BV resistant H1975 and A549 tumors. We found that primary and acquired resistance was associated with distinct patterns of tumor vascularization. In H1975 BV-treated xenografts, MVD was significantly higher at progression compared with short-term treatment ($p<0.01$), returning to levels comparable with vehicle-treated controls (Figure 6A, upper panel – Figure 6B, left panel). In A549 xenografts, there was a significant increase in MVD of BV-treated xenografts at progression compared with controls ($p<0.05$) (Figure 6A, lower panel – Figure 6B, right panel). Taken together, these data suggest that BV

therapy has a dramatic initial anti-angiogenic effect on sensitive H1975 xenografts, but that this effect is lost after continued exposure to the drug, and that therapeutic resistance is associated with revascularization at levels comparable with, or higher than, those in vehicle-treated controls.

Previous results have demonstrated that anti-angiogenic therapy can alter the morphology of the tumor-associated vasculature (23-26). To evaluate the tumor vascularization in greater detail, we assessed vascular tortuosity in vehicle and BV-treated H1975 and A549 xenografts. We found that short-term administration of BV led to a modest, but not statistically significant, reduction in the vessel tortuosity of H1975 tumors (Figure 6A upper panel – Figure 6C, left panel). However, as these tumors developed BV resistance, we noted a 4-fold reduction in vascular tortuosity compared with controls ($p < 0.01$), with larger vessels and a greater degree of pericyte coverage (termed “normalized revascularization”). In contrast, we observed that in A549 xenografts with primary resistance to BV, tumor vascularization was associated with smaller, more tortuous vessels with reduced pericyte coverage compared with controls ($p < 0.05$; Figure 6A, lower panel – Figure 6C, right panel) (termed “sprouting vascularization”), demonstrating that in these models, acquired and primary resistance is associated with different patterns of tumor vascularization.

Dual Blockade of EGFR and VEGFR2 Signaling Pathways Delays Tumor Growth of NSCLC Xenografts

To elucidate whether targeting functioning stromal signaling pathways in BV-resistant tumors can abrogate therapeutic resistance, we targeted EGFR using either the EGFR TKI erlotinib in combination with BV, or the dual VEGFR/EGFR inhibitor vandetanib. Both A549 and H1975 cells are known to be resistant to erlotinib and vandetanib *in vitro*, which is thought to be due to the presence of a KRAS mutation and a secondary EGFR mutation (T790M), respectively (20-22). Consistent with previous results, erlotinib did not inhibit H1975 tumor growth compared with vehicle (Figure 6D), as five (5/6) xenografts progressed, with a median PFS of 12 days. One xenograft progressed while receiving erlotinib plus BV therapy; vandetanib caused tumor growth inhibition in all the xenografts and only two of these progressed. The median PFS in the vandetanib group was 211 days, whereas it was not reached in erlotinib + BV group.

In A549 xenografts, treatment with erlotinib resulted in median PFS values of 53 days, compared with 22 days for vehicle-treated controls ($p=0.12$, log-rank test; Figure 6E). Two tumors progressed on erlotinib plus BV treatment over the course of the experiment (median PFS not reached), and the addition of the EGFR TKI erlotinib to BV significantly delayed the onset of resistance compared with BV alone ($p<0.05$, log-rank test). Only one tumor progressed while receiving vandetanib after 102 days of treatment ($p<0.05$, vandetanib vs. BV). These findings indicate that EGFR inhibition not only reduced the number of NSCLC

xenografts that progressed on therapy compared with BV alone, but also delayed the onset of resistance to VEGF signalling inhibition.

Given the aforementioned EGFR expression in pericytes in the H1975 model, we examined whether targeting EGFR would impact vessel maturation and pericyte coverage. Multicolor IF staining was performed using antibodies directed to CD31 (red) and desmin (green) and pericyte coverage was quantified. In H1975 BV-resistant xenografts, the percentage of blood vessels surrounded by pericytes was 50% higher compared with control tumors ($p < 0.01$). However, pericyte coverage was significantly reduced in tumors receiving long-term treatment with erlotinib plus BV or vandetanib ($p < 0.01$), which likely induced an inhibition of the EGFR pathway activation on the pericytes (Figure 6F). In contrast, A549 xenografts that progressed on BV therapy had significantly fewer blood vessels supported by pericytes compared with controls ($p < 0.05$); nevertheless, long-term administration of erlotinib plus BV or vandetanib also decreased the pericyte coverage in this model compared with controls ($p < 0.05$; see Supplemental Figure 4), providing further support for the role of EGFR in tumor stroma.

Discussion

Early reports examining the effects of anti-angiogenic therapies raised the hopes that these agents may significantly slow or stop tumor growth, and that therapeutic resistance to these agents would be less likely to occur in part because the target was diploid and not prone to the same genetic instability as tumor cells (27, 28). However, both preclinical studies and clinical experience in lung cancer and other solid tumors (13, 15, 29-33) indicate that the vast majority of solid tumors either have primary (de novo) resistance, or will eventually acquire resistance, to the effects of antiangiogenic agents. Although to date most studies of therapeutic resistance to anticancer drugs have focused on the role of tumor cells, recent studies have suggested that host factors, including tumor stroma, may play a role as well (32, 34-38)

In this study, we used mouse- and human-specific profiling of human NSCLC xenografts in mice to investigate stromal and tumor cell alterations occurring in tumors with primary and acquired resistance to BV. This analysis revealed that modulation of angiogenesis-related genes during the development of the resistant phenotype in this model occurs predominantly in the stromal, and not tumor cell, compartment. We validated that two stromal genes, *Egfr* and *Fgfr2* become upregulated during the resistant phase and, moreover, demonstrated that therapeutic inhibition of the EGFR signaling in BV-resistant tumors significantly blocked stromal changes associated with acquired resistance, including increased pericyte coverage, and delayed the onset of resistance.

We studied tumor growth in two human NSCLC xenografts models in mice: H1975, which were initially inhibited by BV but subsequently and sporadically acquired resistance at times ranging from one to six months, and A549, which were not significantly inhibited from the onset of treatment. Human NSCLC cells were used to facilitate species-specific profiling to separately assess tumor and stroma gene expression. These particular models were chosen because of their differences in initial responsiveness to BV and the known resistance of the tumor cells to EGFR blockade (20) which allowed us to assess the impact of stromal EGFR blockade.

We observed that in the acquired resistance model (H1975) BV treatment was initially associated with tumor growth inhibition, a marked reduction in tumor MVD and an increase in tumor endothelial apoptosis after two weeks. These observations are consistent with data from earlier preclinical and clinical reports involving multiple tumor types (23, 24, 26, 39-42). Acquired resistance, however, was associated with a marked increase in overall MVD to levels similar to those of untreated tumors. Furthermore, resistance was not associated with increased phosphoVEGFR2. Consistent with this result, previous reports showed continuing molecular efficacy of the VEGFR2 signaling blockade during the angiogenic evasion phase in a model of pancreatic neuroendocrine tumors (15). Our findings, therefore, provide evidence that VEGFR2-independent mechanisms were contributing to the observed changes in the tumor vasculature. In fact, the vasculature of tumors with acquired resistance had a further, dramatic reduction in vessel tortuosity compared to controls, with greater pericyte

coverage and reduced endothelial apoptosis (“normalized revascularization”). Different patterns of vascular changes were observed in the A549 primary resistance model, with increased MVD at progression associated with greater vessel tortuosity and decreased pericyte coverage (“sprouting vascularization”). Studies have shown that the number of pericytes associated with the tumor vasculature has important clinical implications in that it can determine the efficacy of therapy (43, 44). Pericytes play key roles in maintenance of cancer vasculature by producing survival signaling and may provide a protective mechanism against antiangiogenic therapies. Our findings provide evidence that pericytes may play an important role in the acquisition of resistance in tumors that initially responded to VEGF blockade.

Gene expression analysis showed that among the differentially modulated genes of tumor cells and tumor stroma associated with acquired resistance there is, upregulation of *Egfr* and *Fgfr2*, as well as other members of these signaling pathways, such as *epiregulin (Epgn)*, *Amphiregulin (Areg)*, *Fgf13*, and *Fgfbp1*. A functional gene-interaction network analyses highlighted the EGFR pathway as a major set of interacting genes that were altered in resistant tumors. Alterations in the stromal EGFR and FGFR2 pathways in tumors with acquired resistance were then validated by qRT-PCR using species-specific probes, and further investigation revealed that in resistant tumors, increased activated EGFR was localized almost exclusively on VSCs consistent with pericytes. No significant phospho-EGFR was detectable on VSCs of control tumors. The dual inhibition of VEGFR/EGFR pathways reduced pericyte coverage of tumor vessels and

delayed tumor progression compared with BV alone. A trend towards longer PFS was observed with dual inhibition compared to BV alone as well. Interestingly, primary resistance in A549 xenografts, which was associated with a pattern of “sprouting vascularization”, also had increased phosphoEGFR immunoreactivity compared with controls but in this model it was localized almost exclusively to tumor endothelium, not VSCs. As expected, dual VEGFR/EGFR inhibition did not reduce pericyte coverage in this model, but did significantly delay the emergence of resistance compared with BV alone. These findings demonstrate that increases in stromal EGFR signaling are associated with tumor progression in two different models of BV resistance and that targeting the EGFR pathway is able to reverse changes associated with resistance (e.g. pericyte coverage) and delay the emergence of resistance. It also highlights that a signaling pathway may play different roles in tumor stroma depending on the cellular context.

Studies examining the EGFR distribution on endothelium suggest that it is restricted to blood vessels supplying pathologic tissues (45), where it activates angiogenic programs (46). Others have reported that EGFR is activated on endothelium when tumor cells express EGFR ligands, such as transforming growth factor alpha (TGF α) or EGF (47, 48). We localized p-EGFR to the blood vessels of primary resistant A549 xenografts and to the vascular supporting pericytes of H1975 xenografts that acquired BV-resistance. Consistent with this observation, a recent study found that the EGFR TKI gefitinib significantly suppressed tumor-associated pericyte function (49). We also found that EGFR blockade decreased pericyte coverage in this model. To our knowledge, this is

the first evidence demonstrating a potential role for EGFR signaling in pericytes or other stromal cells in resistance to VEGF pathway inhibitors or other antiangiogenic agents.

A recent study has identified a role for PDGF-C expressed by tumor associated fibroblasts in VEGF inhibitor resistance (38) and in attenuating tumor response to anti-VEGF treatment in a model of glioblastoma (50). Somewhat surprisingly, we did not observe upregulation of any PDGF receptors (PDGFR) or ligands. In contrast, we observed a modest but statistically significant *downregulation* of stromal genes encoding PDGF-A, PDGF-B and PDGFR- β . Given the role of the PDGF family in multiple tumor processes, including pericyte recruitment and function (51, 52), it appears that pericyte-expressed EGFR may play a complimentary or compensatory role in the increased pericyte coverage observed in the acquired resistance model.

Activation of FGF/FGFR pathway has been shown to be a critical regulator of the “angiogenic switch” (53) and to be upregulated in response to anti-angiogenic therapy (15). We observed an approximately 6-fold increase in stromal *Fgfr2* gene expression in tumors with acquired resistance and, consistent with this finding, an increase in the number of FGFR2-expressing cells in these tumors. This immunoreactivity appeared to be largely, but not exclusively, on tumor endothelium. This suggests that the FGFR2 pathway may promote VEGF-independent endothelial survival as previously observed in other preclinical models (54, 55) although we cannot rule out the possibility that it plays a role in other non-endothelial stromal cells. Circulating levels of the FGFR2 ligand bFGF

were also elevated in the plasma of mice with BV-resistant tumors. This observation is notable in light of our recent observation that acquired resistance to chemotherapy and BV in colorectal cancer patients is associated with an increase in circulating bFGF (14), suggesting that similar mechanisms may be occurring in cancer patients.

Other genes differentially expressed in tumors with acquired resistance included *Thrombospondin 2* (*Thbs2*), a potent negative regulator of angiogenesis (56) and *Bax*, a pro-apoptotic member of the Bcl-2 family of proteins (57), which were both downregulated in resistant tumors. Recent studies have shown Amphiregulin to be involved in resistance to targeted therapy in NSCLC cells by promoting BAX inhibition (58). Consistent with these observations, our results showed an upregulation of *Areg* gene accompanied by downregulation of *Bax* in the stromal compartment of BV resistant tumors compared with controls.

The mechanisms underlying regulation of tumor stromal genes altered in resistant tumors remains to be established and are likely to differ in the various stromal cell types. Expression of many of the genes, including *FGFR2* (59) and *EGFR* family members and *thrombospondin-2* (as well as tumor cell *carbonic anhydrase-9*), are known to be regulated by hypoxia or to correlate with expression of hypoxia inducible factor 1 alpha (HIF1 α), as reviewed in (60). One possible explanation is that BV therapy initially triggers a significant decrease in tumor MVD and an increase in tumor hypoxia (61), triggering upregulation of hypoxia-dependent pathways. It is worth noting, however, that BV resistance was not associated with significant increases in many stromal genes known to be

upregulated by hypoxia, and many of the genes upregulated in BV resistance are not known to be hypoxia-regulated. Hypoxia is likely, therefore, to be only one of many factors- both host- and tumor cell dependent- that are likely to be impacting the tumor and its microenvironment in resistant tumors.

To date, studies of therapeutic resistance have almost exclusively focused on mechanisms involving tumor cells. Our findings provide evidence that there are substantial changes in tumor stroma associated with resistance in our models, and in fact a far greater number of stromal genes than tumor cell genes were significantly different in tumors with acquired resistance compared with vehicle controls. Pathway analysis implicated EGFR as a central pathway upregulated in resistant stroma. EGFR localization differed in the two models: in pericytes of highly “normalized” vessels (acquired resistance) and tumor endothelium (primary resistance). As expected, the addition of EGFR blockade reduced pericyte coverage in the acquired resistance model, and delayed the emergence of resistance.

Murine xenograft models such as those used in this study have several limitations, such as absence of an intact immune system, and it is likely that different types of tumors, or tumors growing in locations (e.g. lung), may employ different mechanisms. Nevertheless, taken as a whole, these findings have important scientific and clinical implications. They demonstrate that stromal mechanisms can play a role- perhaps, in some cases a dominant role- in impacting the responsiveness of a tumor. In this regard, it is worth noting that stromal cells typically comprise a minority of cells in a tumor, and that detecting

stromal-specific changes associated with resistance (for example, endothelial phosphoEGFR) requires methodologies different than those in typical clinical use. The results also highlight a previously unappreciated role for stromal EGFR in therapeutic resistance, and illustrate that EGFR may play different roles in stroma depending on the cellular context. Effects on stromal cells may therefore contribute to the clinical benefit observed in patients treated with EGFR inhibitors even when tumors were negative for EGFR by IHC (62) or FISH (63), or did not contain activating mutations (64). Finally, they suggest that approaches targeting stromal resistance pathways may enhance the efficacy of regimens containing VEGF inhibitors. It is worth noting, however, that while combinations of VEGF and EGFR pathway inhibition have shown promise in NSCLC (65, 66), therapeutic resistance nevertheless continues to emerge, indicating that additional resistance mechanisms remain to be uncovered.

Materials and Methods

***In Vivo* Studies**

All animal studies reported were approved from the local committee for animal care. To generate tumor xenografts, A549 and H1975 tumor cells (2.0×10^6) were injected in 100 μ l of HBSS into the subcutaneous flank of 4 to 8 weeks old male athymic nude mice (NCI-nu). Body weights and tumor volumes were recorded twice weekly. Tumor volumes (mm^3) were calculated as $= \pi/6(a)^2(b)$, where a is the smaller measurement of the tumor and b is the larger one. When the tumor volumes reached an average of approximately 270 mm^3 , mice were

randomly assigned to one of six treatment groups: (a) control intraperitoneal (i.p.) injection of vehicle (PBS) twice weekly; (b) i.p. injection of BV (10 mg/kg) twice weekly; (c) control, oral (p.o.) administration of vehicle daily; (d) erlotinib (100 mg/kg) p.o. daily; (e) erlotinib p.o. daily plus BV i.p. twice weekly; (f) vandetanib (50 mg/kg) p.o. daily. Tumors were sacrificed due to tumor burden. Tumors were excised and one portion was fixed in formalin and embedded in paraffin; another portion was embedded in OCT (Miles, Inc., Elkhart, IN) and rapidly frozen in liquid nitrogen. Additional tumor sections for molecular studies were snap-frozen in liquid nitrogen. Staining with hematoxylin and eosin (H&E) was used to confirm the presence of tumor in each sample included in the analysis. For short-term treatment studies, tumor-bearing animals were treated for 2 weeks with vehicle and BV and then sacrificed. Tumor tissues were collected for immunohistochemical studies.

RNA Microarray Analysis

Total RNA was extracted from snap-frozen tissues using the *mirVana*[™] miRNA Isolation Kit (Ambion, Austin, TX) and 1 µg of total RNA was amplified using the Illumina[®] TotalPrep RNA Amplification Kit (Ambion) according to the manufacturer's instructions. Labeled cRNA was hybridized on mouseWG-6 v2 and human WG-6v3 Expression BeadChips (Illumina[®], San Diego, CA) for analysis of murine and human transcriptomes. Signal intensities determined by streptavidin-Cy3 fluorescence were scanned with a Sherlock_1000 Array Scanner (Ambion). Data were analyzed using the BRB-ArrayTools Version 3.7.0

Beta platform developed by Dr. Richard Simon [<http://linus.nci.nih.gov/BRB-ArrayTools.html>]. A log base 2 transformation was applied to the data set prior to data normalization. A median array was selected as the reference array for normalization and statistical significance was set using a $p < 0.01$. To evaluate the expression of genes involved in response to hypoxia, lymphangiogenesis and angiogenesis in BV-resistant xenografts compared with controls, a gene list of 269 genes used in previous publications and associated with and involved in these processes and was compared (67). Genes differentially expressed between groups were determined applying univariate t-test with estimation of the false discovery rate (FDR). Genes were determined using selection criteria of a $p < 0.005$ and a fold-change ≥ 1.5 .

Determination of Microvessel Density (MVD), vessel tortuosity and pericyte coverage of tumor-associated blood vessels

Tumor MVD was determined as previously described (68). In brief, we examined tumors microscopically to identify hot spots by low magnification (x100), and the mean MVD was quantified as the total number of CD31⁺ structures observed in a minimum of five microscopic fields at higher power of vision per tumor (x200). For each group, tumors from four different mice receiving short- and long-term treatment were used. As previously described (69), the tortuosity (T) of blood vessel was defined as $T = (L/S) - 1$, where L is the length of the vessel of interest and S is the straight-line distance between its endpoints. The vessel length (L)

was evaluated in 4 samples for each treatment group by tracing along the midline of the blood vessels that showed up in a longitudinal cut within an image (100x) and the number of pixels was converted into distance in millimeters with Image J v1.34 software (NIH).

To determine the extent of pericyte coverage on the tumor vasculature, tumor sections were stained for CD31 (red) and desmin (green) as described above. Five fields in each tumor were randomly identified at original magnification x200, and those blood vessels that were at least 50% covered by green-desmin positive cells were considered to be positive for pericyte coverage.

Standard Methods

Reagents, tumor cell lines and conditions, and standard techniques for qRT-PCR, IHC, confocal microscopy and LSC are described in the Supplemental Data.

Plasma bFGF Concentration Analysis

Basic FGF levels were measured in the plasma of tumor-bearing animals by multiplex bead assay (BioRad, Hercules, CA; Millipore, Billerica, MA) in a 96-well plate according to the manufacturer's protocol. The concentrations were calculated based on a standard curve derived by performing six serial dilutions of

a protein standard in assay diluent. Plasma samples were tested in duplicate and the mean value used for analysis.

Statistics

Statistical significance was tested using GraphPad Prism 5 software (GraphPad Software Inc., San Diego, CA). For comparison between two groups, Student's t test and log-rank test were used. A p value < 0.05 on two-tailed testing was considered significant.

Supplemental Data

The Supplemental Data include Supplemental Experimental Procedures, 2 supplemental tables and 4 supplemental figures.

ACKNOWLEDGMENTS

This study was supported in part by the Department of Defense BATTLE W81XWH-06-1-0303, Department of Defense PROSPECT award W81XWH-07-1-03060, the University of Texas Southwestern Medical Center, The University of Texas M. D. Anderson Cancer Center Lung SPORE NIH grant P50 CA070907, and research support from AstraZeneca Cross Alliance. J.V. Heymach is a Damon Runyon-Lilly Clinical Investigator supported in part by Damon Runyon

Cancer Research Foundation grant CI 24-04 and is also supported by the Physician Scientist Program at M. D. Anderson Cancer Center. We gratefully acknowledge Dr. Michael Worley for editing the manuscript, Drs. Anderson Ryan, Juliane Jürgensmeier, and Lee Ellis for helpful scientific discussions, and AstraZeneca (Macclesfield, United Kingdom) for providing vandetanib. Research in the laboratory of Fortunato Ciardiello is supported by a grant from the Associazione Italiana per la Ricerca sul Cancro (AIRC).

References

1. Folkman, J., and Shing, Y. 1992. Angiogenesis. *J Biol Chem* 267:10931-10934.
2. Folkman, J. 1971. Tumor angiogenesis: therapeutic implications. *N Engl J Med* 285:1182-1186.
3. Carmeliet, P., and Jain, R.K. 2000. Angiogenesis in cancer and other diseases. *Nature* 407:249-257.
4. Hanahan, D., and Folkman, J. 1996. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* 86:353-364.
5. Ferrara, N., Gerber, H.P., and LeCouter, J. 2003. The biology of VEGF and its receptors. *Nat Med* 9:669-676.
6. Ferrara, N., and Davis-Smyth, T. 1997. The biology of vascular endothelial growth factor. *Endocr Rev* 18:4-25.
7. Ellis, L.M., and Hicklin, D.J. 2008. VEGF-targeted therapy: mechanisms of anti-tumour activity. *Nat Rev Cancer* 8:579-591.
8. Kerbel, R.S. 2008. Tumor angiogenesis. *N Engl J Med* 358:2039-2049.
9. Jain, R.K., di Tomaso, E., Duda, D.G., Loeffler, J.S., Sorensen, A.G., and Batchelor, T.T. 2007. Angiogenesis in brain tumours. *Nat Rev Neurosci* 8:610-622.
10. Ferrara, N., Hillan, K.J., Gerber, H.P., and Novotny, W. 2004. Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. *Nat Rev Drug Discov* 3:391-400.
11. Dvorak, H.F. 2002. Vascular permeability factor/vascular endothelial growth factor: a critical cytokine in tumor angiogenesis and a potential target for diagnosis and therapy. *J Clin Oncol* 20:4368-4380.
12. Nilsson, M., and Heymach, J.V. 2006. Vascular endothelial growth factor (VEGF) pathway. *J Thorac Oncol* 1:768-770.
13. Sandler, A., Gray, R., Perry, M.C., Brahmer, J., Schiller, J.H., Dowlati, A., Lilenbaum, R., and Johnson, D.H. 2006. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N Engl J Med* 355:2542-2550.
14. Kopetz, S., Hoff, P.M., Morris, J.S., Wolff, R.A., Eng, C., Glover, K.Y., Adinin, R., Overman, M.J., Valero, V., Wen, S., et al. Phase II Trial of Infusional Fluorouracil, Irinotecan, and Bevacizumab for Metastatic Colorectal Cancer: Efficacy and Circulating Angiogenic Biomarkers Associated With Therapeutic Resistance. *J Clin Oncol* 28:453-459.
15. Casanovas, O., Hicklin, D.J., Bergers, G., and Hanahan, D. 2005. Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors. *Cancer Cell* 8:299-309.
16. Yu, J.L., Rak, J.W., Coomber, B.L., Hicklin, D.J., and Kerbel, R.S. 2002. Effect of p53 status on tumor response to antiangiogenic therapy. *Science* 295:1526-1528.
17. Davis, D.W., Takamori, R., Raut, C.P., Xiong, H.Q., Herbst, R.S., Stadler, W.M., Heymach, J.V., Demetri, G.D., Rashid, A., Shen, Y., et al. 2005. Pharmacodynamic analysis of target inhibition and endothelial cell death in tumors treated with the vascular endothelial growth factor receptor antagonists SU5416 or SU6668. *Clin Cancer Res* 11:678-689.

18. Heymach, J.V., Desai, J., Manola, J., Davis, D.W., McConkey, D.J., Harmon, D., Ryan, D.P., Goss, G., Quigley, T., Van Den Abbeele, A.D., et al. 2004. Phase II Study of the Antiangiogenic Agent SU5416 in Patients with Advanced Soft Tissue Sarcomas. *Clin Cancer Res* 10:5732-5740.
19. Paez-Ribes, M., Allen, E., Hudock, J., Takeda, T., Okuyama, H., Vinals, F., Inoue, M., Bergers, G., Hanahan, D., and Casanovas, O. 2009. Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell* 15:220-231.
20. Naumov, G.N., Nilsson, M.B., Cascone, T., Briggs, A., Straume, O., Akslen, L.A., Lifshits, E., Byers, L.A., Xu, L., Wu, H.K., et al. 2009. Combined Vascular Endothelial Growth Factor Receptor and Epidermal Growth Factor Receptor (EGFR) Blockade Inhibits Tumor Growth in Xenograft Models of EGFR Inhibitor Resistance. *Clin Cancer Res* 15:3484-3494.
21. Kobayashi, S., Boggon, T.J., Dayaram, T., Janne, P.A., Kocher, O., Meyerson, M., Johnson, B.E., Eck, M.J., Tenen, D.G., and Halmos, B. 2005. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 352:786-792.
22. Pao, W., Wang, T.Y., Riely, G.J., Miller, V.A., Pan, Q., Ladanyi, M., Zakowski, M.F., Heelan, R.T., Kris, M.G., and Varmus, H.E. 2005. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2:e17.
23. Jain, R.K. 2005. Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. *Science* 307:58-62.
24. Batchelor, T.T., Sorensen, A.G., di Tomaso, E., Zhang, W.T., Duda, D.G., Cohen, K.S., Kozak, K.R., Cahill, D.P., Chen, P.J., Zhu, M., et al. 2007. AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. *Cancer Cell* 11:83-95.
25. Greenberg, J.I., and Cheresh, D.A. 2009. VEGF as an inhibitor of tumor vessel maturation: implications for cancer therapy. *Expert Opin Biol Ther* 9:1347-1356.
26. Jain, R.K. 2001. Normalizing tumor vasculature with anti-angiogenic therapy: A new paradigm for combination therapy. *Nat Med* 7:987-989.
27. Boehm, T., Folkman, J., Browder, T., and O'Reilly, M.S. 1997. Antiangiogenic therapy of experimental cancer does not induce acquired drug resistance. *Nature* 390:404-407.
28. Kerbel, R.S. 1991. Inhibition of tumor angiogenesis as a strategy to circumvent acquired resistance to anti-cancer therapeutic agents. *Bioessays* 13:31-36.
29. Heymach, J.V., Paz-Ares, L., De Braud, F., Sebastian, M., Stewart, D.J., Eberhardt, W.E., Ranade, A.A., Cohen, G., Trigo, J.M., Sandler, A.B., et al. 2008. Randomized phase II study of vandetanib alone or with paclitaxel and carboplatin as first-line treatment for advanced non-small-cell lung cancer. *J Clin Oncol* 26:5407-5415.
30. Broxterman, H.J., Lankelma, J., and Hoekman, K. 2003. Resistance to cytotoxic and anti-angiogenic anticancer agents: similarities and differences. *Drug Resist Updat* 6:111-127.
31. Kerbel, R.S., Yu, J., Tran, J., Man, S., Vilorio-Petit, A., Klement, G., Coomber, B.L., and Rak, J. 2001. Possible mechanisms of acquired resistance to anti-

- angiogenic drugs: implications for the use of combination therapy approaches. *Cancer Metastasis Rev* 20:79-86.
32. Ellis, L.M., and Hicklin, D.J. 2008. Pathways mediating resistance to vascular endothelial growth factor-targeted therapy. *Clin Cancer Res* 14:6371-6375.
 33. Heymach, J.V., Sledge, G.W., and Jain, R.K. 2010. Tumor Angiogenesis. In *Holland-Frei Cancer Medicine* 8. W.K. Hong, R.C. Bast Jr., W.N. Hait, D.W. Kufe, R.E. Pollock, R.R. Weichselbaum, J.F. Holland, and E. Frei III, editors. Hamilton: BC Decker.
 34. Ferrara, N. 2009. Pathways mediating VEGF-independent tumor angiogenesis. *Cytokine Growth Factor Rev*.
 35. Crawford, Y., and Ferrara, N. 2009. Tumor and stromal pathways mediating refractoriness/resistance to anti-angiogenic therapies. *Trends Pharmacol Sci* 30:624-630.
 36. Ebos, J.M., Lee, C.R., and Kerbel, R.S. 2009. Tumor and host-mediated pathways of resistance and disease progression in response to antiangiogenic therapy. *Clin Cancer Res* 15:5020-5025.
 37. Shojaei, F., Wu, X., Malik, A.K., Zhong, C., Baldwin, M.E., Schanz, S., Fuh, G., Gerber, H.P., and Ferrara, N. 2007. Tumor refractoriness to anti-VEGF treatment is mediated by CD11b+Gr1+ myeloid cells. *Nat Biotechnol* 25:911-920.
 38. Crawford, Y., Kasman, I., Yu, L., Zhong, C., Wu, X., Modrusan, Z., Kaminker, J., and Ferrara, N. 2009. PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment. *Cancer Cell* 15:21-34.
 39. Roland, C.L., Dineen, S.P., Lynn, K.D., Sullivan, L.A., Dellinger, M.T., Sadegh, L., Sullivan, J.P., Shames, D.S., and Brekken, R.A. 2009. Inhibition of vascular endothelial growth factor reduces angiogenesis and modulates immune cell infiltration of orthotopic breast cancer xenografts. *Mol Cancer Ther* 8:1761-1771.
 40. Kim, K.J., Li, B., Winer, J., Armanini, M., Gillett, N., Phillips, H.S., and Ferrara, N. 1993. Inhibition of vascular endothelial growth factor-induced angiogenesis suppresses tumour growth in vivo. *Nature* 362:841-844.
 41. Willett, C.G., Boucher, Y., di Tomaso, E., Duda, D.G., Munn, L.L., Tong, R.T., Chung, D.C., Sahani, D.V., Kalva, S.P., Kozin, S.V., et al. 2004. Direct evidence that the VEGF-specific antibody bevacizumab has antivascular effects in human rectal cancer. *Nat Med* 10:145-147.
 42. Tong, R.T., Boucher, Y., Kozin, S.V., Winkler, F., Hicklin, D.J., and Jain, R.K. 2004. Vascular normalization by vascular endothelial growth factor receptor 2 blockade induces a pressure gradient across the vasculature and improves drug penetration in tumors. *Cancer Res* 64:3731-3736.
 43. Lu, C., Thaker, P.H., Lin, Y.G., Spannuth, W., Landen, C.N., Merritt, W.M., Jennings, N.B., Langley, R.R., Gershenson, D.M., Yancopoulos, G.D., et al. 2008. Impact of vessel maturation on antiangiogenic therapy in ovarian cancer. *Am J Obstet Gynecol* 198:477 e471-479; discussion 477 e479-410.
 44. Bergers, G., Song, S., Meyer-Morse, N., Bergsland, E., and Hanahan, D. 2003. Benefits of targeting both pericytes and endothelial cells in the tumor vasculature with kinase inhibitors. *J Clin Invest* 111:1287-1295.

45. Amin, D.N., Hida, K., Bielenberg, D.R., and Klagsbrun, M. 2006. Tumor endothelial cells express epidermal growth factor receptor (EGFR) but not ErbB3 and are responsive to EGF and to EGFR kinase inhibitors. *Cancer Res* 66:2173-2180.
46. Cheng, H., Langley, R.R., Wu, Q., Wu, W., Feng, J., Tsan, R., Fan, D., and Fidler, I.J. 2005. Construction of a novel constitutively active chimeric EGFR to identify new targets for therapy. *Neoplasia* 7:1065-1072.
47. Wu, W., O'Reilly, M.S., Langley, R.R., Tsan, R.Z., Baker, C.H., Bekele, N., Tang, X.M., Onn, A., Fidler, I.J., and Herbst, R.S. 2007. Expression of epidermal growth factor (EGF)/transforming growth factor- α by human lung cancer cells determines their response to EGF receptor tyrosine kinase inhibition in the lungs of mice. *Mol Cancer Ther* 6:2652-2663.
48. Kuwai, T., Nakamura, T., Sasaki, T., Kim, S.J., Fan, D., Villares, G.J., Zigler, M., Wang, H., Bar-Eli, M., Kerbel, R.S., et al. 2008. Phosphorylated epidermal growth factor receptor on tumor-associated endothelial cells is a primary target for therapy with tyrosine kinase inhibitors. *Neoplasia* 10:489-500.
49. Iivanainen, E., Lantta, S., Zhang, N., Tvorogov, D., Kulmala, J., Grenman, R., Salven, P., and Elenius, K. 2009. The EGFR inhibitor gefitinib suppresses recruitment of pericytes and bone marrow-derived perivascular cells into tumor vessels. *Microvasc Res* 78:278-285.
50. di Tomaso, E., London, N., Fuja, D., Logie, J., Tyrrell, J.A., Kamoun, W., Munn, L.L., and Jain, R.K. 2009. PDGF-C induces maturation of blood vessels in a model of glioblastoma and attenuates the response to anti-VEGF treatment. *PLoS One* 4:e5123.
51. Pietras, K., Pahler, J., Bergers, G., and Hanahan, D. 2008. Functions of paracrine PDGF signaling in the proangiogenic tumor stroma revealed by pharmacological targeting. *PLoS Med* 5:e19.
52. Ostman, A., and Heldin, C.H. 2007. PDGF receptors as targets in tumor treatment. *Adv Cancer Res* 97:247-274.
53. Kandel, J., Bossy-Wetzel, E., Radvanyi, F., Klagsbrun, M., Folkman, J., and Hanahan, D. 1991. Neovascularization is associated with a switch to the export of bFGF in the multistep development of fibrosarcoma. *Cell* 66:1095-1104.
54. Paris, F., Fuks, Z., Kang, A., Capodieci, P., Juan, G., Ehleiter, D., Haimovitz-Friedman, A., Cordon-Cardo, C., and Kolesnick, R. 2001. Endothelial apoptosis as the primary lesion initiating intestinal radiation damage in mice. *Science* 293:293-297.
55. Karsan, A., Yee, E., Poirier, G.G., Zhou, P., Craig, R., and Harlan, J.M. 1997. Fibroblast growth factor-2 inhibits endothelial cell apoptosis by Bcl-2-dependent and independent mechanisms. *Am J Pathol* 151:1775-1784.
56. Zhang, X., and Lawler, J. 2007. Thrombospondin-based antiangiogenic therapy. *Microvasc Res* 74:90-99.
57. Wei, M.C., Zong, W.X., Cheng, E.H., Lindsten, T., Panoutsakopoulou, V., Ross, A.J., Roth, K.A., MacGregor, G.R., Thompson, C.B., and Korsmeyer, S.J. 2001. Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. *Science* 292:727-730.

58. Busser, B., Sancey, L., Jossierand, V., Niang, C., Favrot, M.C., Coll, J.L., and Hurbin, A. 2009. Amphiregulin Promotes BAX Inhibition and Resistance to Gefitinib in Non-small-cell Lung Cancers. *Mol Ther*.
59. Giatromanolaki, A., Koukourakis, M.I., Sivridis, E., Turley, H., Talks, K., Pezzella, F., Gatter, K.C., and Harris, A.L. 2001. Relation of hypoxia inducible factor 1 alpha and 2 alpha in operable non-small cell lung cancer to angiogenic/molecular profile of tumours and survival. *Br J Cancer* 85:881-890.
60. Semenza, G.L. 2003. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* 3:721-732.
61. Franco, M., Man, S., Chen, L., Emmenegger, U., Shaked, Y., Cheung, A.M., Brown, A.S., Hicklin, D.J., Foster, F.S., and Kerbel, R.S. 2006. Targeted anti-vascular endothelial growth factor receptor-2 therapy leads to short-term and long-term impairment of vascular function and increase in tumor hypoxia. *Cancer Res* 66:3639-3648.
62. Chung, K.Y., Shia, J., Kemeny, N.E., Shah, M., Schwartz, G.K., Tse, A., Hamilton, A., Pan, D., Schrag, D., Schwartz, L., et al. 2005. Cetuximab shows activity in colorectal cancer patients with tumors that do not express the epidermal growth factor receptor by immunohistochemistry. *J Clin Oncol* 23:1803-1810.
63. Italiano, A., Follana, P., Caroli, F.X., Badetti, J.L., Benchimol, D., Garnier, G., Gugenheim, J., Haudebourg, J., Keslair, F., Lesbats, G., et al. 2008. Cetuximab shows activity in colorectal cancer patients with tumors for which FISH analysis does not detect an increase in EGFR gene copy number. *Ann Surg Oncol* 15:649-654.
64. Tsao, M.S., Sakurada, A., Cutz, J.C., Zhu, C.Q., Kamel-Reid, S., Squire, J., Lorimer, I., Zhang, T., Liu, N., Daneshmand, M., et al. 2005. Erlotinib in lung cancer - molecular and clinical predictors of outcome. *N Engl J Med* 353:133-144.
65. Natale, R.B., Bodkin, D., Govindan, R., Sleekman, B.G., Rizvi, N.A., Capo, A., Germonpre, P., Eberhardt, W.E., Stockman, P.K., Kennedy, S.J., et al. 2009. Vandetanib Versus Gefitinib in Patients With Advanced Non-Small-Cell Lung Cancer: Results From a Two-Part, Double-Blind, Randomized Phase II Study. *J Clin Oncol* 27:2523-2529.
66. Herbst, R.S., O'Neill, V.J., Fehrenbacher, L., Belani, C.P., Bonomi, P.D., Hart, L., Melnyk, O., Ramies, D., Lin, M., and Sandler, A. 2007. Phase II study of efficacy and safety of bevacizumab in combination with chemotherapy or erlotinib compared with chemotherapy alone for treatment of recurrent or refractory non small-cell lung cancer. *J Clin Oncol* 25:4743-4750.
67. Van den Eynden, G.G., Van Laere, S.J., Van der Auwera, I., Gilles, L., Burn, J.L., Colpaert, C., van Dam, P., Van Marck, E.A., Dirix, L.Y., and Vermeulen, P.B. 2007. Differential expression of hypoxia and (lymph)angiogenesis-related genes at different metastatic sites in breast cancer. *Clin Exp Metastasis* 24:13-23.
68. Weidner, N., Semple, J.P., Welch, W.R., and Folkman, J. 1991. Tumor angiogenesis and metastasis--correlation in invasive breast carcinoma. *N Engl J Med* 324:1-8.
69. Stockmann, C., Doedens, A., Weidemann, A., Zhang, N., Takeda, N., Greenberg, J.I., Cheresch, D.A., and Johnson, R.S. 2008. Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis. *Nature* 456:814-818.

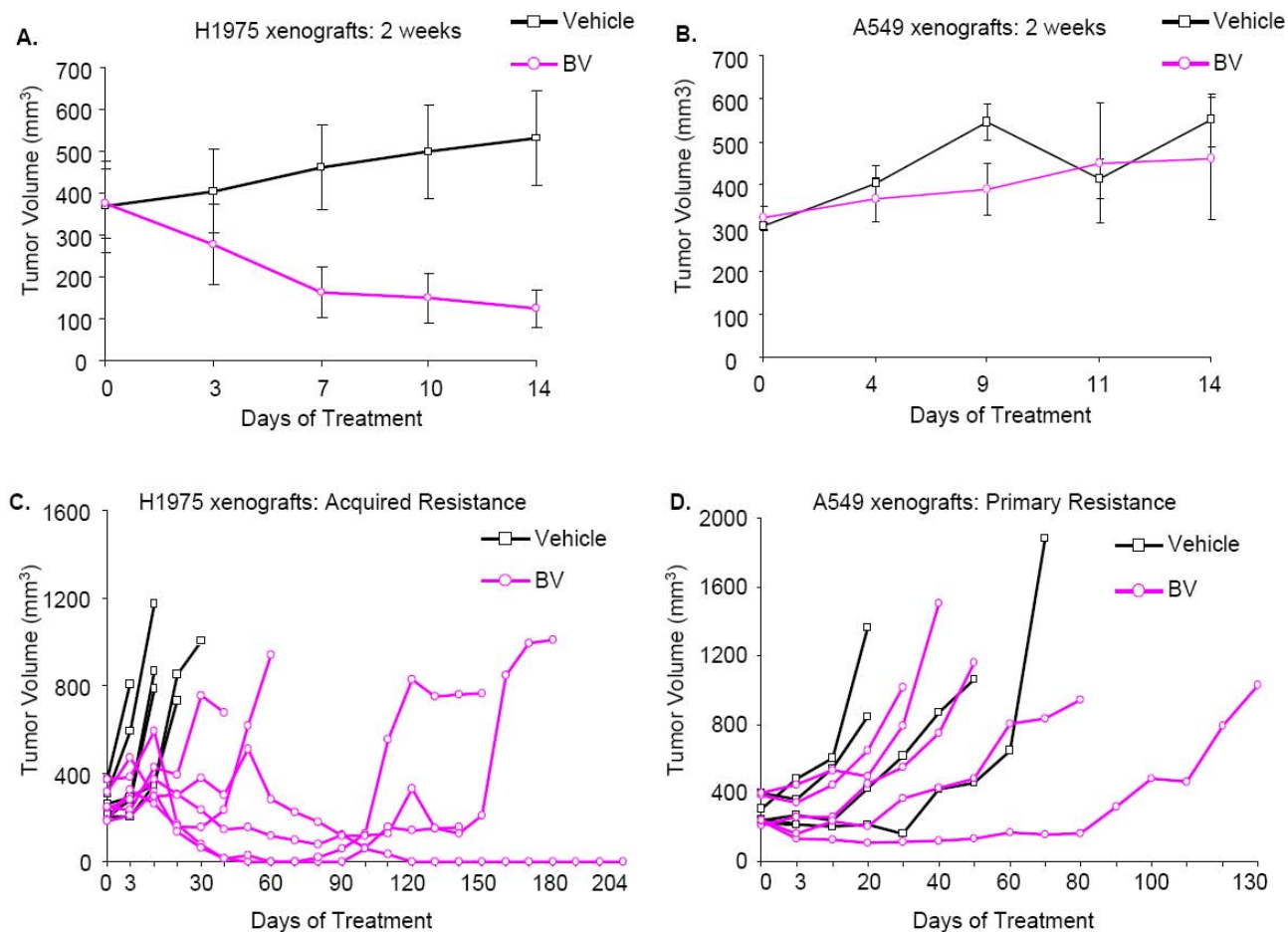


Figure 1. H1975 and A549 NSCLC Xenografts Show Acquired and Primary Resistance to BV

A, B. Tumor growth curves of H1975 xenografts receiving vehicle (control) and BV (n=5) for 2 weeks (short-term treatment).

C, D. Individual tumor growth curves of H1975 (A) and A549 (B) vehicle- and BV-treated xenografts. Vehicle- and BV-treated tumors are shown in black and pink, respectively. The median PFS in H1975 BV-resistant xenografts was 138 days compared with 7 days in the control group ($p < 0.001$ BV vs. vehicle, log-rank test); in A549 BV-resistant tumors the median PFS was 35 days compared with 22 days in controls ($p = 0.16$ BV vs. vehicle, log-rank test).

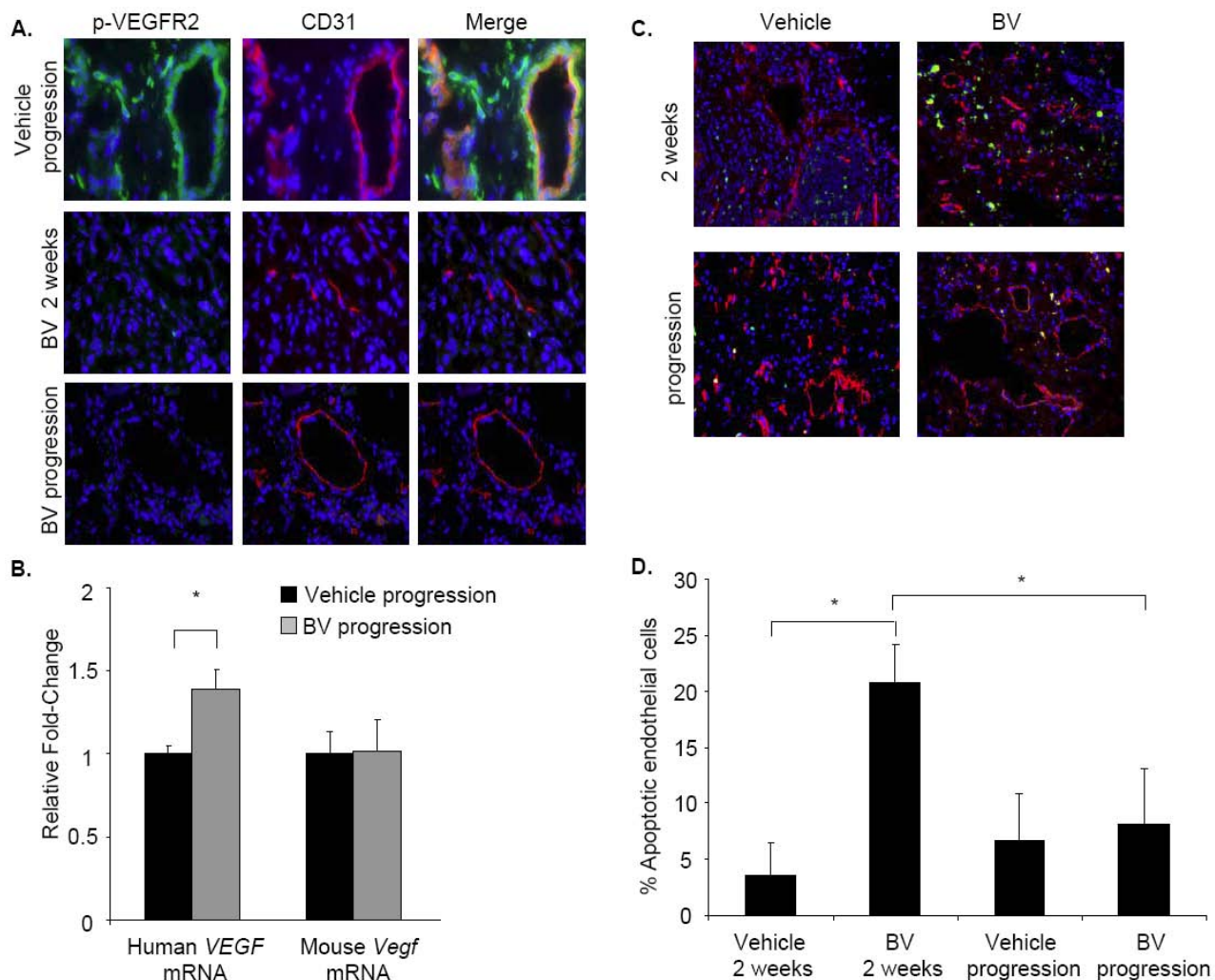


Figure 2. BV Resistance Is Associated with Sustained p-VEGFR2 Inhibition and Decreased Endothelial Cell Apoptosis

A. Representative IF staining of H1975 vehicle and BV-treated tumors for 2 weeks and at progression show p-VEGFR-2 (green) and CD31⁺ cells (red). Blue shows the nuclei.

B. Human and mouse *VEGFA* mRNA expression in H1975 vehicle- and BV-treated tumors at progression using qRT-PCR (n=4, each group). Human *GAPDH* and mouse tubulin were used as housekeeping controls for human and mouse mRNA, respectively. Data were normalized relative to mRNA levels in vehicle-progression samples and are graphed as relative fold change \pm SEM, *p<0.05 (t-test).

C. Representative immunofluorescent images of CD31 (red) and TUNEL (green) staining in H1975 xenografts after 2 weeks and at progression with vehicle and BV. Blue shows the nuclei.

D. Percentage of apoptotic ECs quantified by IF staining, labeling the endothelial marker CD31 and using the terminal TUNEL kit to identify apoptotic cells. The percentage of CD31⁺ and TUNEL⁺ cells was counted in a minimum of 5 microscopic fields (200 \times) in each of minimum 3 tumor samples per each group. Data are graphed as the percentage \pm SEM, *p<0.05 (t-test).

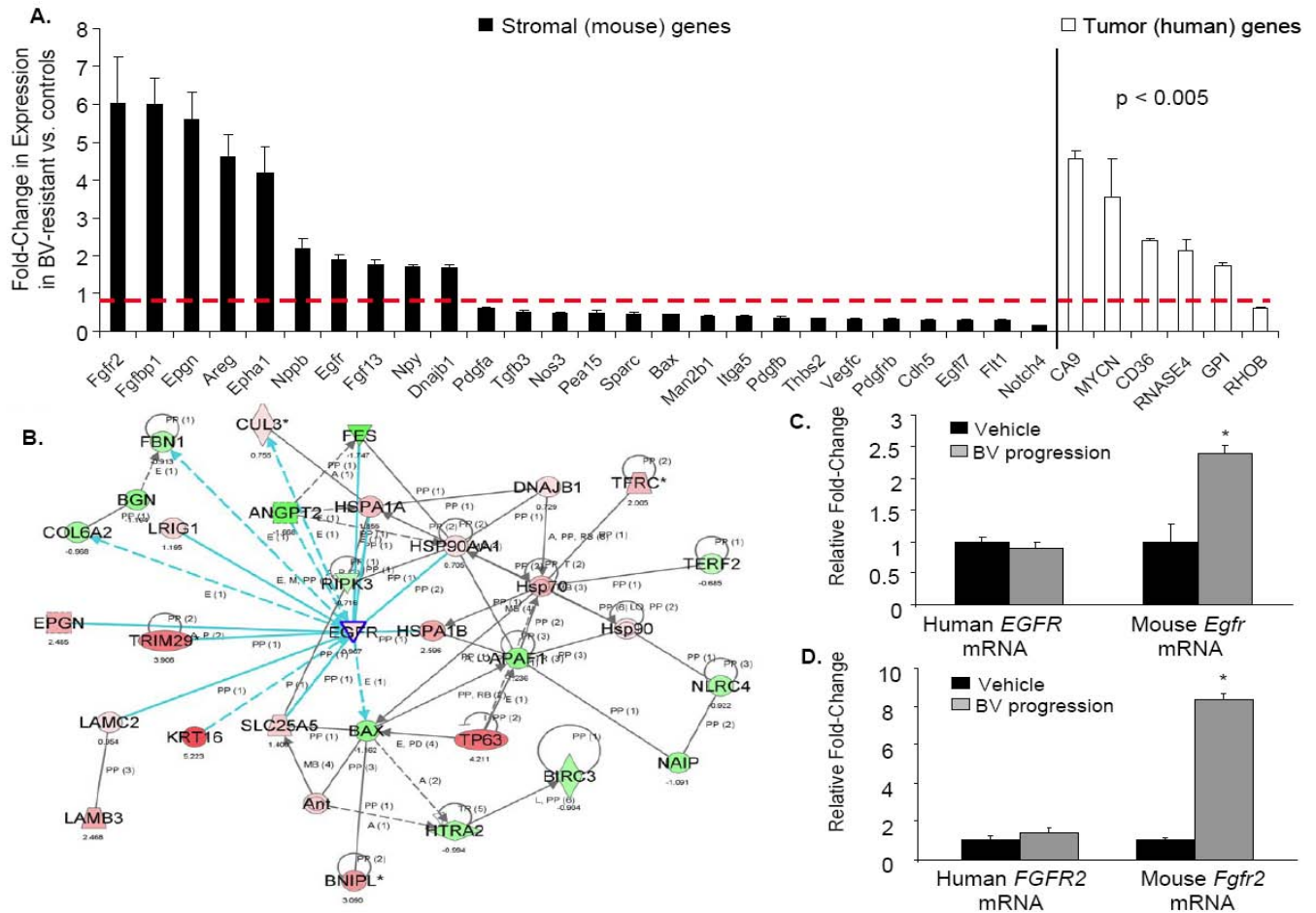


Figure 3. BV Resistance Is Associated with Increased Expression of Stromal Genes Involved in Angiogenesis

A. Stromal and human angiogenic genes differentially regulated in H1975 BV-resistant xenografts compared with vehicle-controls (n=3, each group). Two-sample t-test ($p < 0.005$) with random variance model was applied. Exact permutation p-values for significant genes were computed based on 10 available permutations. Data are graphed as differences in fold change \pm SEM of genes in BV-resistant tumors vs. controls. The dashed red line defines the fold change of gene expression at 1, indicating no changes in the expression of a specific gene in BV-resistant tumors vs. controls. **B.** Functional pathway analysis of selected genes and their interaction nodes in a gene network significantly modulated between the BV-resistant and control xenografts mouse stroma. The network score was calculated by the inverse log of the p-value and indicates the likelihood of focus genes in a network being found together than due to chance. The selected genes (*Egfr*, *Bax* and *Dnabp1*) and their interaction segments are highlighted by a blue border. Gene expression variation by at least 1.5-fold is depicted by color (red, upregulated; green, down-regulated; grey, no significant change). **C, D.** Human and mouse *EGFR* (C) and *FGFR2* (D) mRNA expression in H1975 vehicle- and BV- progression xenografts (n=4, each group) using qRT-PCR. Human *GAPDH* and mouse tubulin were used as housekeeping controls for human and mouse mRNA, respectively. Data are normalized relative to mRNA levels in vehicle-progression samples and graphed as relative fold change \pm SEM, * $p < 0.05$ (t-test).

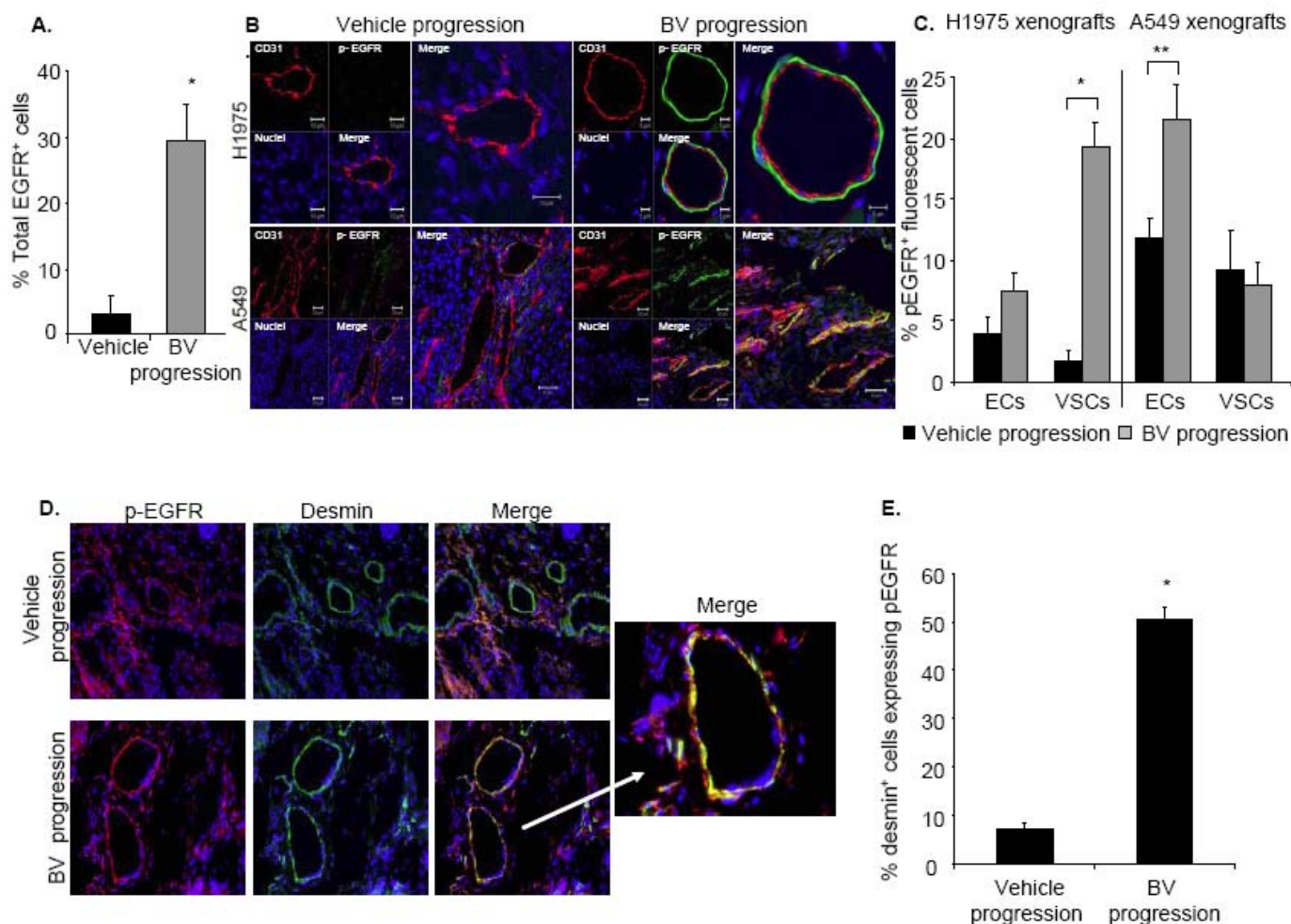


Figure 4. BV Resistance Is Associated with Increased EGF R Activation on the Vascular Supporting Cells and the Tumor Vasculature

A. Quantification of EGFR-expressing cells in H1975 tumors following vehicle and BV treatment at progression (n=4, each group), using LCS. Data are graphed as the mean \pm SEM, *p<0.01 (t-test). **B.** Representative IF staining of CD31 (red) and p-EGFR (green) using confocal microscopy in vehicle- and BV-treated H1975 (top panel) and A549 xenografts (lower panel) at progression (n=4, each group in both xenograft models). **C.** Quantification of vascular supporting cells (VSCs) and endothelial cells (CD31⁺, ECs) expressing p-EGFR in H1975 and A549 vehicle- and BV-treated tumors at progression (n=4, each group in both the xenograft models). Phospho-EGFR⁺ cells were counted in a minimum of 5 random microscopic fields for each tumor sample at 200 \times magnification. Data are graphed as percentage \pm SEM, *p<0.01, **p<0.05 (t-test). **D.** Representative immunofluorescent images of p-EGFR (red) and desmin (green) positive staining in H1975 vehicle- and BV-treated H1975 xenografts at progression. The white arrow shows the overlapping of p-EGFR and desmin in BV-resistant H1975 tumors at higher magnification (400 \times). **E.** Quantification of desmin⁺ cells expressing p-EGFR in H1975 vehicle- and BV-treated H1975 tumors at progression. Phospho-EGFR⁺ cells were counted in a minimum of 5 random microscopic fields for each tumor sample at 200 \times (n=4, each group). Data are graphed as percentage \pm SEM, *p<0.01, (t-test).

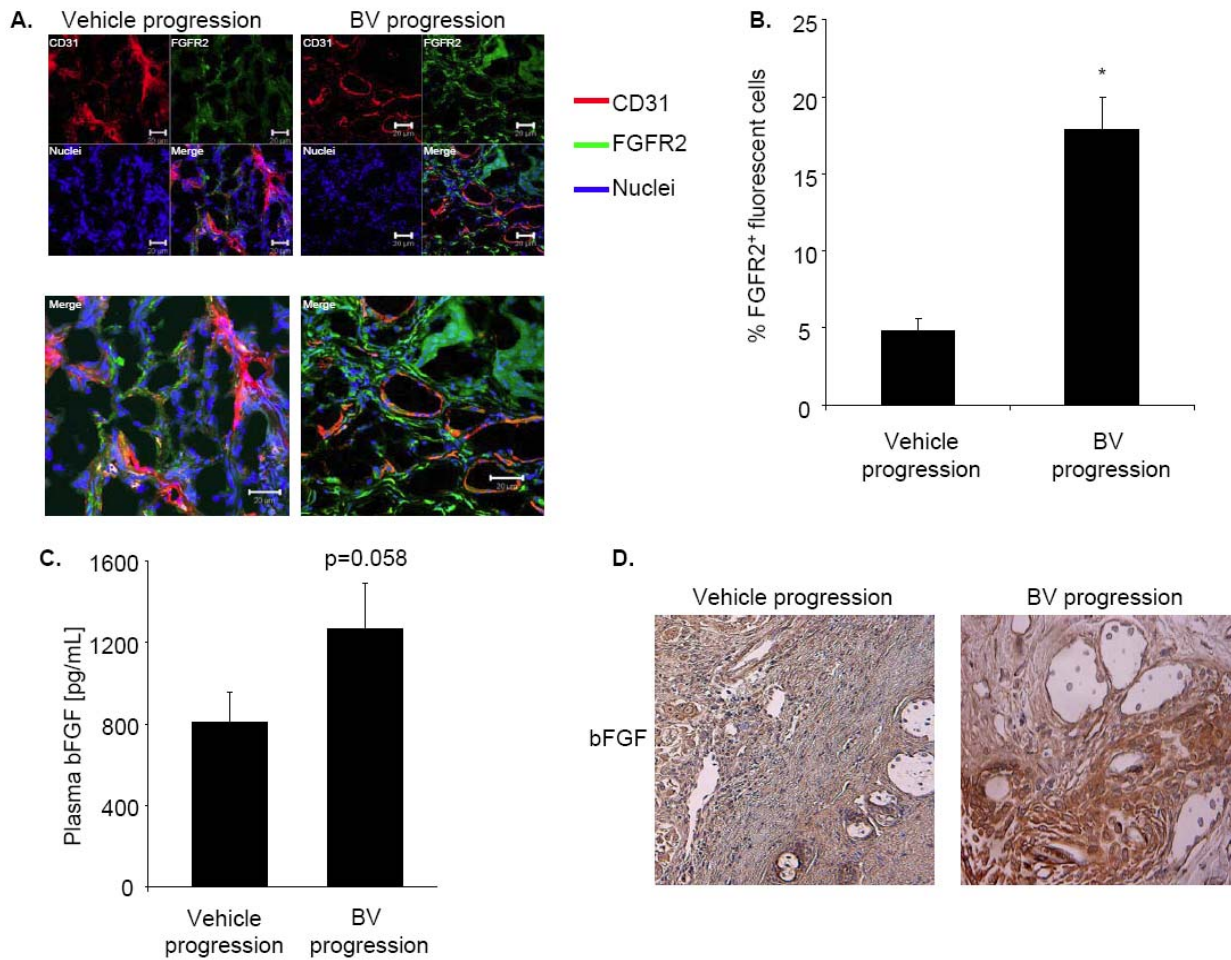


Figure 5. Increase in Stromal FGFR2 Expression in H1975 BV-Resistant Xenografts

A. Representative immunofluorescent images of CD31 (red) and FGFR2 (green) staining in H1975 vehicle- and BV-treated H1975 xenografts at progression, using confocal microscopy (200 \times).

B. Quantification of FGFR2⁺ cells (green) counted in 5 random microscopic fields (200 \times ; n=4, each group). Data are graphed as the percentage \pm SEM, *p<0.01 (t-test).

C. Basic FGF (bFGF) levels (pg/mL) were measured in the plasma of vehicle- and BV-treated H1975 xenografts at progression, by multiplex bead assay. Data are graphed as the mean \pm SEM, p=0.058 (t-test).

D. Representative IHC staining of bFGF in vehicle- and BV-treated H1975 tumors at time of progression (n=4, per each group).

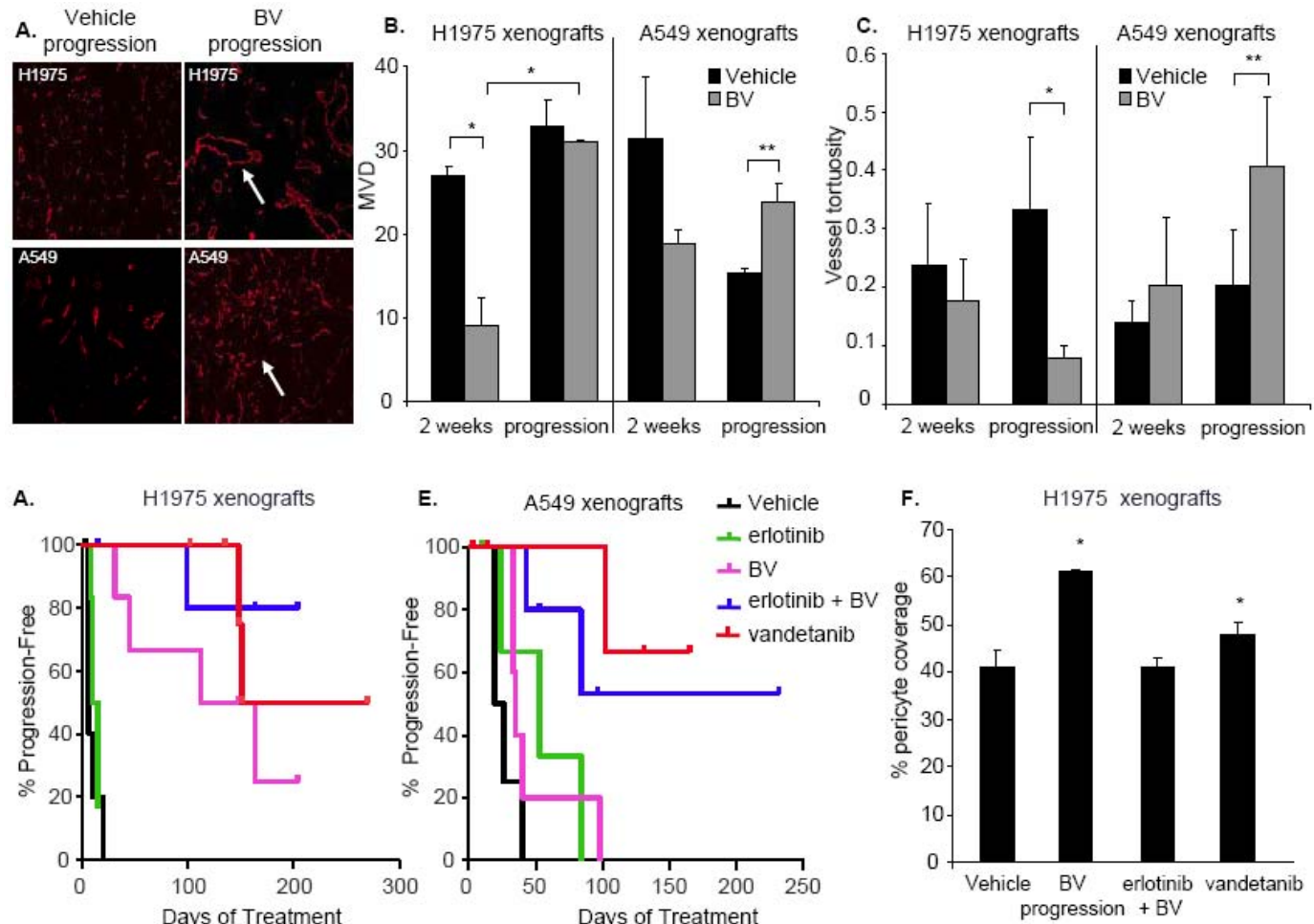


Figure 6. Altered patterns of tumor vascular density, tortuosity, and pericyte coverage in BV-resistant tumors, and impact of combined EGFR/VEGFR inhibition

A. CD31⁺ staining (red) showing representative images of vasculature in H1975 (upper panel) and A549 (lower panel) tumors after short-term (2 weeks) treatment with BV and at progression.

B, C. Quantification of MVD (B) and tumor vessel tortuosity (C) in the same tumors. CD31⁺ vessels were counted in 5 microscopic fields in each of 4 samples per each group at 200 \times , shown as mean \pm SEM, * p<0.01, ** p<0.05 (t-test).

D, E. Kaplan-Meier plots of PFS in H1975 (D) and A549 xenografts (E). H1975 xenografts: BV vs. vehicle p<0.01; BV vs. erlotinib p<0.01; erlotinib + BV vs. BV p=0.24, erlotinib + BV vs. vehicle p<0.01; vandetanib vs. BV p=0.295; vandetanib vs. vehicle p<0.01 (log-rank test). A549 xenografts: BV vs. vehicle p = 0.16; BV vs. erlotinib p = 0.99; erlotinib + BV vs. BV p = 0.044; erlotinib + BV vs. vehicle p<0.01 vandetanib vs. BV p = 0.012; vandetanib vs. vehicle p = 0.015 (log-rank test).

F. Pericyte coverage in H1975 xenografts. The percentage of CD31⁺ vessels with at least 50% coverage of associated desmin⁺ cells was counted in 5 microscopic fields at 200 \times in tumors that progressed on the indicated treatment (n=4, vehicle and BV groups; n=2, vandetanib group, and n=1, erlotinib + BV). Data are graphed as percentage \pm SEM, *p<0.05 (t-test).

High Expression of Folate Receptor Alpha in Lung Cancer Correlates with Adenocarcinoma Histology and *EGFR* Mutation

Maria Ines Nunez¹, Carmen Behrens², Denise M. Woods¹, Heather Lin³, Milind Suraokar², Humam Kadara², John D. Minna^{5,6}, Wayne Hofstetter⁴, Neda Kalhor¹, Waun K. Hong², J. Jack Lee³, Wilbur Franklin⁷, David J. Stewart², and Ignacio I. Wistuba^{1,2*}

Departments of ¹Pathology, ²Thoracic/Head and Neck Medical Oncology, ³Biostatistics, and ⁴Thoracic and Cardiovascular Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, TX; ⁵The Hamon Center for Therapeutic Oncology Research and ⁶Departments of Internal Medicine and Pharmacology, The University of Texas Southwestern Medical Center, Dallas, TX; and, ⁷Department of Pathology, University of Colorado at Denver and Health Sciences Center, Aurora, CO

Running title: Expression of FR α and RFC1 in NSCLC

Type of manuscript: Research article / **Section:** Prognostic and predictive factors

Acknowledgments of research support: This study was supported in part by grants from the Department of Defense (W81XWH-07-1-0306 to J.D.M., W.K.H., D.J.S., and I.I.W.), the Specialized Program of Research Excellence in Lung Cancer (P50CA70907 to J.D.M. and I.I.W.), and the National Cancer Institute (Cancer Center Support Grant CA-16672).

Key Words: non-small cell lung carcinoma, EGFR, membrane transporter, FR α , RFC1

Abbreviations: EGFR, epidermal growth factor receptor; FFPE, formalin-fixed paraffin-embedded; FR α , folate receptor alpha; RFC1, reduced folate carrier 1; TS, thymidylate synthase; NSCLC, non-small cell lung carcinoma; OR, odds ratio.

***Corresponding author:** Ignacio I. Wistuba, MD, Department of Pathology, Box 085, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 54, Houston, Texas 77030, Email: iwistuba@mdanderson.org; Phone: (713) 563-9184; Fax: (713) 792-0309

Statement of Translational Relevance.

Membrane transporters FR α and RFC1 are potential biomarkers of tumor response to antifolate chemotherapy. Information on the protein expression of these receptors in NSCLC is limited. Here, we report for the first time that NSCLC frequently overexpressed FR α and RFC1 proteins by studying a large series of cases with annotated clinico-pathologic information. Importantly, we report that tumor cells from lung adenocarcinoma histology expressed significantly higher levels of cytoplasmic and membrane FR α than squamous cell carcinoma, and tumors from never-smokers were significantly more likely to express cytoplasmic FR α than those from smokers. In lung adenocarcinomas, the presence of *EGFR* mutations correlated with higher expression of membrane FR α and *FOLR1* gene expression. We postulate that this information may be useful in selecting which patients with NSCLC may benefit from and should receive treatment with antifolate inhibiting agents, including pemetrexed.

ABSTRACT

Purpose. Folate receptor alpha (FR α) and receptor folate carrier-1 (RFC1) regulate cellular uptake of folate molecules inside the cell, and are potential biomarkers of tumor response to antifolate chemotherapy. Information on the protein expression of these receptors in non-small cell lung carcinoma (NSCLC) is limited.

Experimental Design. Expressions of FR α and RFC1 were examined by immunohistochemistry in 320 surgically resected NSCLC (202 adenocarcinomas and 118 squamous cell carcinomas) tissue specimens using tissue microarrays. The findings were correlated with patient clinicopathologic characteristics. *FOLR1* mRNA expression was examined using publicly available microarray datasets. FR α expression was correlated with tumors' thymidylate synthase (TS) and p53 expression in NSCLCs, and with *EGFR* and *KRAS* mutations in adenocarcinomas.

Results. NSCLC frequently overexpressed both FR α and RFC1. In a multivariate analysis, lung adenocarcinomas were more likely to express FR α in the cytoplasm (odds ratio [OR] = 4.39; $P < 0.0001$) and membrane (OR = 5.34; $P < 0.0001$) of malignant cells than squamous cell carcinomas. Tumors from never-smokers were significantly more likely to express cytoplasmic (OR = 3.35; $P < 0.03$) and membrane (OR = 3.60; $P = 0.0005$) FR α than those from smokers. FR α and RFC1 expressions did not correlate with NSCLC patients' outcome. In adenocarcinoma, *EGFR* mutations correlated with higher expression of membrane FR α and *FOLR1* gene expressions. FR α and TS expression inversely and significantly correlated ($P = 0.03$).

Conclusions. Membrane transporter FR α and RFC1 proteins are frequently overexpressed in NSCLC tumor tissues. The higher levels of FR α in lung adenocarcinomas than in squamous cell carcinomas may explain these tumors' different responses to antifolate chemotherapy.

INTRODUCTION

Despite intensive research on molecular targeted therapy, chemotherapy still represents the main treatment option for patients with advanced non-small cell lung cancer (NSCLC; stages IIIB and IV) (1). Over recent years chemotherapy after surgical resection has become the standard of care for treatment of selected patients with early stage (i.e., stage IB, II, or IIIA) NSCLC (2). However, most tumors develop drug resistance, leading to chemotherapy failure. The factors associated with chemotherapy resistance are not well understood, but some phenomena have been associated with this resistance, including, among others, decreases or alterations in the membrane transporters involved in drug uptake systems or increase in drug efflux pumps (3).

In advanced NSCLC, doublet combinations of a platinum drug with another chemotherapy agent are the reference regimens (1). A recent Phase III trial showed that cisplatin/gemcitabine offered a significantly longer survival compared to cisplatin/pemetrexed in patients with squamous cell carcinoma (SCC) (4). These findings prompted restriction of the pemetrexed regimen to patients with NSCLC of “non-squamous” histology only (5). Pemetrexed is a potent inhibitor of thymidylate synthase (TS) (6, 7) and other folate-dependent enzymes, including dihydrofolate reductase and glycinamide ribonucleotide formyl transferase (8). The mechanisms involved in the varying outcomes of patients treated with pemetrexed based on tumor histologic type are unknown.

Folic acid and its reduced congeners are required for one carbon transfer reactions that are used in the synthesis of nucleotide bases, amino acids, and other methylated compounds, and consequently, they are required in larger quantities by proliferating cells (3). FR α is a glycoprotein that is anchored to the apical cell membrane surface of a variety of polarized normal epithelial cells (9) and binds folate at a high affinity to mediate transport into the cytoplasm of cells (3). RFC1 is more ubiquitously expressed in normal cells, binds folate at low affinity, and represents the sole folate uptake pathway for most normal cells (10).

In this study we aimed to characterize the expression of FR α and RFC1 proteins in a large series of surgically resected NSCLC tissue specimens with annotated clinicopathologic features. We correlated the expression of FR α with the expression of TS. Our findings of higher expression of FR α expression in lung tumors with adenocarcinoma histology and tumors obtained from never-smokers prompted us to correlate the expression of FR α with tumors' epidermal growth factor receptor (*EGFR*) and *KRAS* mutation status in adenocarcinomas, and with tumors' p53 protein expression in all NSCLCs.

MATERIALS AND METHODS

Case selection and tissue microarray (TMA) construction. We obtained archived formalin-fixed and paraffin-embedded (FFPE) NSCLC tissues from the Lung Cancer Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX). These lung cancer specimens were resected with curative intent between 1997 and 2001. The institutional review board at M. D. Anderson Cancer Center approved our study. The tissue specimens were histologically classified according to the 2004 World Health Organization classification (11). We constructed TMAs using three 1-mm diameter cores. Detailed clinico-pathologic information was available for most cases (Table 1).

Immunohistochemical staining and evaluation. To test the expression of the membrane transporters we used a monoclonal homemade antibody against FR α (clone Mb343, IgG), dilution 1:500 (12), and a polyclonal antibody against RFC1 (Abcam, Cambridge, MA), dilution 1:100. To assess the expression of TS, we used a monoclonal antibody (Zymed Carlsbad, CA, USA), dilution 1:100. For p53 analysis, we used mouse monoclonal antihuman p53, clone DO7 (Dako, Carpinteria, CA), dilution 1:400.

For FR α we used a previously published immunohistochemistry protocol (12). For RFC1 and TS, immunohistochemical staining was performed as follows: 5- μ M FFPE tissue sections were deparaffinized and hydrated, and underwent heat-induced epitope retrieval in a DAKO antigen retrieval bath at 121°C for 30 seconds and 90°C for 10 seconds in a decloaking chamber (Biocare, Concord, CA), followed by a 30-min cool down. Prior to antibody immunostaining, endogenous peroxidase activity was blocked with 3% H₂O₂ in methanol for 30 min. To block nonspecific antibody binding, tissue sections were incubated with 10% fetal bovine serum in Tris-buffered saline solution with Tween 20 for 30 min. The slides were incubated with primary antibody at ambient temperature for 60 min for all antibodies. This was

followed by incubation with biotin-labeled secondary antibody (Envision Dual Link +, DAKO) for 30 min. Staining was developed with 0.05% 3',3'-diaminobenzidine tetrahydrochloride, which had been freshly prepared in 0.05 mol/L Tris buffer at pH 7.6 containing 0.024% H₂O₂, and then the slides were counterstained with hematoxylin, dehydrated, and mounted.

Two observers (M.N. and I.W.) jointly quantified the immunohistochemical expression of the membrane transporters (magnification 20×) in normal bronchial epithelium and lung tumor malignant epithelial cells. For each membrane transporter and TS, we defined 3 categories of intensity of immunostaining (0 to 3+). Next, an expression score (range, 0–300) was obtained by multiplying the intensity of staining by the percent of cells (0-100%) staining. p53 expression was categorized by percentage of tumor cells expressing nuclear p53 as positive (≥5%) or negative (0-5%).

***EGFR* and *KRAS* mutation analysis.** Exons 18 through 21 of *EGFR* and exon 1 of *KRAS* were amplified by polymerase chain reaction (PCR) using intron-based primers as previously described (13, 14).

Assessment of membrane transporter expression in microarray datasets. The cancer microarray database and integrated data-mining platform Oncomine (15) was utilized to analyze the expression of *FOLR1* (FR α) and *SLC19A1* (RFC1), and in microarray databases of NSCLC available online (16-19). The statistical significances of differences in expression of the genes were provided by Oncomine and confirmed by a two-tailed t-test with random variance. Gene expression data of lung adenocarcinomas with annotated mutation data of *EGFR* and *KRAS* were obtained from the Ladanyi and Gerald laboratories at the Memorial Sloan-Kettering Cancer Center (MSKCC) (http://cbio.mskcc.org/Public/lung_array_data/) (20). Available Affymetrix® raw data files of the transcriptomes of 190 adenocarcinomas (set I, n=88; set II,

n=102) were analyzed using the BRB-ArrayTools version 3.7.0 software developed by using the BRB-ArrayTools v.3.7.0 developed by Dr. Richard Simon and BRB-ArrayTools Development Team (21). Robust multi-array analysis (RMA) was used for normalization of gene expression data using the R language environment (22). *FOLR1* mRNA expression levels in both MSKCC datasets were median-centered by the Cluster v.2.11 software. Differences in normalized median-centered *FOLR1* expression levels were assessed for statistical significance by the two-tailed test and $P < 0.05$ were considered statistically significant.

Statistical methods. Associations between biomarker expression scores and patient clinico-pathologic data were assessed using the Wilcoxon's rank sum test or Kruskal-Wallis test, as appropriate, for continuous variables and the chi square test for categorical variables. For recurrence free survival (RFS) and overall survival (OS) analyses, we tested binary cutoff points of biomarkers using the median expression score for each marker. Univariate and multivariate Cox proportional hazards regression models were used to assess the effects of covariates on survival. All statistical tests were two-sided, and P values <0.05 were considered statistically significant.

RESULTS

Immunohistochemical expression of FR α and RFC1 in NSCLC. Both adenocarcinoma and SCC expressed relatively high levels of FR α and RFC1 in the malignant cells (Fig. 1 and Table 2). For FR α , the average expression scores and frequency of any expression (score >0) were significantly higher in adenocarcinomas than in SCCs at membrane ($P<0.001$) and cytoplasmic ($P<0.001$) localizations (Fig. 2). Both NSCLC histologies demonstrated similar levels of cytoplasmic and membrane RFC1 expression. For both markers the tumor cells exhibited stronger immunohistochemical expression than the 11 samples of normal bronchial epithelia adjacent to tumors (data not shown).

Correlation of FR α and RFC1 expression with clinicopathologic features. The multivariate analysis of the immunohistochemical expression of the two membrane transporters as a dichotomized variable (positive, score >0, vs. negative, score =0), after adjustment for patient's tumor histology, smoking history, sex, and disease stage, revealed that adenocarcinomas were more likely than SCCs to express cytoplasmic (odds ratio [OR] = 4.39; $P<0.0001$) and membrane (OR = 5.34; $P<0.0001$) FR α . In addition, tumors from never-smokers were significantly more likely to express cytoplasmic (OR = 3.35; $P < 0.03$) and membrane (OR = 3.60; $P=0.0005$) FR α than those of smokers. In the multivariate analysis, the patient's sex was not an independent significant factor influencing tumor expression of FR α . No correlation was found between expression of both membrane transporters and RFS or OS in 230 patients with stage I or II NSCLCs (median follow up, 7.2 years).

Correlation between FR α expression and tumors' p53 expression and EGFR and KRAS mutation status. Our findings of higher expression of FR α expression in lung tumors with adenocarcinoma histology and tumors obtained from never-smokers prompted us to correlate

the expression of FR α with tumors' *EGFR* and *KRAS* mutation status in adenocarcinomas, and with tumors' p53 protein expression in all NSCLCs.

In lung adenocarcinomas, *EGFR* mutant tumors demonstrated significantly higher expression scores for membrane FR α (mean scores: mutant 134.8 vs. wild-type 67.1; $P=0.002$) than wild-type tumors. No correlation between FR α expression and adenocarcinoma tumors' *KRAS* mutation status was detected.

Of all NSCLCs tested, 38% (75/195) of adenocarcinomas and 69% (80/116) of SCCs had a positive p53 level ($\geq 5\%$). Interestingly, we found that the scores for FR α expression in both membrane ($P=0.001$) and cytoplasm ($P<0.001$) were significantly lower in malignant cells from NSCLC tumors with positive p53 expression (mean score: membrane 33.4, SD 59.9, and cytoplasm 58.3, SD 60.0) than in tumors with negative p53 expression (mean score: membrane 65.3, SD 90.6, and cytoplasm 83.55.3, SD 65.3).

***FOLR1* mRNA expression in tumor tissues.** Our findings that protein levels of FR α was greater in adenocarcinomas than in SCCs incited us to analyze expression levels of the mRNA of the *FOLR1* in published microarray datasets of NSCLC tumor specimens and compare them by histologic type (16-19). In accordance with our immunohistochemistry data, *FOLR1* mRNA expression levels were significantly higher in adenocarcinomas ($n=197$) than in SCCs ($n=210$) in all four datasets available: 1.8 vs. 1.0 ($P<0.0001$) (17), 0.81 vs. 0.73 ($P=0.03$) (19), 2.61 vs. 0.98 ($P<0.0001$) (18), and 0.93 vs. 0.31 ($P<0.0001$) (16).

To confirm our findings on the increased FR α immunoreactivity in tumors obtained from *EGFR* mutant lung adenocarcinomas compared to wild type tumors, we probed this association using the mRNA expression levels of *FOLR1* in publicly available microarray datasets with information on *EGFR* and *KRAS* mutation status (20). Notably, the analysis of the microarray data further revealed the statistically significant up-regulation of *FOLR1* mRNA levels in *EGFR*

mutant lung adenocarcinomas compared to wild-type tumors in both available datasets ($P=0.00016$ and $P=0.003$) (Fig. 3). In addition, no statistically significant differences were found in *FOLR1* expression levels between *KRAS* mutant lung adenocarcinomas and wild-type tumors (data not shown). These findings confirm the close positive association between *FOLR1* expression and *EGFR* mutation status which we had found at the protein level by assessment of FR α immunoreactivity.

Correlation of immunohistochemical expression of TS and FR α . TS was expressed frequently in the nucleus and cytoplasm of malignant NSCLC cells. However, the frequency of any TS expression (score >0) was higher in the cytoplasm (212/267, 79%) than in the nucleus (117/267, 44%) of these cells. Although cytoplasmic expression of TS was similar in both NSCLC histologic types (Table 2), nuclear expression was significantly higher ($P=0.003$) in SCCs (mean score: 13.8, SD 27.7) than in adenocarcinomas (mean score: 9.3, SD 27.1). The level of TS expression did not correlate with clinicopathologic characteristics, including RFS and OS. In all NSCLC, significantly ($P=0.02$) higher expression of nuclear TS immunostaining was detected in tumors with positive p53 expression (67/114, 58%) than in those with negative p53 staining (65/147, 44%). In adenocarcinomas, there was no correlation between TS expression and *EGFR* or *KRAS* mutation status.

We correlated the expression of TS and FR α in NSCLC tissue specimens. The score for nuclear TS expression correlated negatively with the score for cytoplasmic FR α expression in SCCs ($r = -0.20$; $P=0.04$), and showed marginally significant negative correlation with membrane FR α expression in adenocarcinomas ($r = -0.16$; $P=0.05$). When we examined the correlation of any expression (score >0) of both markers in tumors, we found that in SCCs expression of nuclear TS was significantly inversely correlated ($P=0.03$) with membrane expression of FR α , and that most tumors positive for TS (62/79, 79%) lacked membrane FR α .

(Table 3). This correlation was not detected in adenocarcinomas.

We determined the pattern of any expression (score >0) of membrane FR α and nuclear TS in both NSCLC tumor types. The most common expression pattern in adenocarcinomas was FR α -positive/TS-negative (68/155, 44%), whereas the most common expression pattern in SCCs was FR α -negative/TS-positive (43/99, 43%) (Table 3).

DISCUSSION

Membrane transporters FR α and RFC1 are potential biomarkers of tumor response to antifolate chemotherapy. Information on the protein expression of these receptors in NSCLC is limited. Here, we report for the first time that NSCLC frequently overexpress FR α and RFC1 proteins by studying a large series of cases with annotated clinico-pathologic information. Importantly, we report that tumor cells from lung adenocarcinoma histology expressed significantly higher levels of cytoplasmic and membrane FR α than squamous cell carcinoma, and tumors from never-smokers were significantly more likely to express cytoplasmic FR α than those from smokers. In lung adenocarcinomas, the presence of *EGFR* mutations correlated with higher expression of membrane FR α and *FOLR1* gene expression. We postulate that this information may be useful in selecting which patients with NSCLC may benefit from and should receive treatment with antifolate inhibiting agents, including pemetrexed.

In lung cancer, the factors and mechanisms associated with chemotherapy resistance are not well understood. In the past several years, several studies have shown that the mechanism associated with resistance to common chemotherapy drugs might be mediated by two broad mechanisms: failure of a sufficient amount of drug to reach the target, and failure to achieve cell death after the drug reaches the target (23). The underlying complex mechanisms by which chemotherapy agents enter into cells remain poorly defined. It has been established recently that, among other factors, some cell membrane transporters may play important roles in chemotherapy drug influx (24, 25).

Our study showed that RFC1 is expressed frequently in malignant cells of NSCLC tumor tissues. The only report available on the expression of RFC1 in human tumors showed relatively high levels of mRNA gene expression in NSCLC, with similar expression in adenocarcinomas and SCCs (26). These data are consistent with our protein expression data showing that levels of expression of RFC1 were similar in the two histologic types.

Interestingly, in our study the expression of membrane and cytoplasmic FR α was significantly higher in adenocarcinomas than in SCCs. FR α has been shown by immunohistochemical studies to be overexpressed in several epithelial tumors (26-30). However, to the best of our knowledge, there is not published report of FR α protein expression in NSCLC tumors and correlation with clinical and pathological features. Our protein expression findings agree with the significantly higher levels of expression of *FLOR1* (FR α gene) mRNA in adenocarcinomas than in SCCs in all four public microarray datasets available (16-19). Similar findings have been reported in a quantitative (q)PCR study of mRNA expression of 119 NSCLC tissue specimens (26).

The findings of higher levels of FR α protein and *FLOR1* mRNA expression in adenocarcinomas than in SCCs of the lung may have important clinical implications. The higher level of FR α protein expression in adenocarcinoma cells may explain the better response of advanced NSCLC of nonsquamous histology when treated with the combination of cisplatin and the multitargeted antifolate agent pemetrexed (4). However, this needs to be further tested in NSCLC tumor tissue specimens obtained from patients treated with pemetrexed. In addition, FR α is currently considered an attractive target for biologic therapy in tumors in which it is overexpressed such as ovarian cancer (31), by using FR α -specific approaches and inducing inhibition of growth under folate-limiting conditions to boost immunity to tumors (32).

Our finding that NSCLCs of never-smokers have a higher expression of FR α than those of smokers is of interest. Our data showing significantly higher cytoplasmic and membrane FR α expression in NSCLCs obtained from never-smokers are in agreement with the previous report of higher levels of mRNA *FLOR1* by qPCR in adenocarcinomas from nonsmokers and light smokers than in those from heavy smokers (26). These differences in the expression of FR α by smoking status are consistent with our findings of higher FR α expression in NSCLCs lacking p53 expression and in adenocarcinomas harbouring *EGFR* mutation, two features associated

with the pathogenesis of non-smoking-related lung cancer (33). Of interest, the analysis of the publicly available microarray data confirmed at mRNA gene expression level our observation that *EGFR* mutant adenocarcinoma tumors expressed higher levels of FR α protein. There are not data available on the response to antifolate chemotherapy agents in lung adenocarcinomas based on *EGFR* mutation status. However, it has been shown that advanced stage adenocarcinoma harbouring this mutation showed improved response to other type of (carboplatin and paclitaxel) chemotherapy (34).

Because of their roles in metabolism of the chemotherapy agent pemetrexed (5, 37), we correlated the expressions of TS and FR α in NSCLC tissue specimens by histologic type. As previously reported (38, 39), TS protein was expressed frequently in the nucleus (44%) and cytoplasm (79%) of malignant NSCLC cells. In our analysis we determined that nuclear expression was significantly higher in SCCs than in adenocarcinomas. Ceppi et al (38) previously reported that immunohistochemical expression of TS mRNA and protein was significantly higher in SCCs of the lung than in adenocarcinomas. In this previously reported immunohistochemical analysis, however, expression of TS in the malignant cells was not distinguished as nuclear or cytoplasmic. It has been shown that low levels of TS mRNA expression significantly correlated with *in vitro* chemosensitivity of freshly explanted human tumor specimens to pemetrexed (40). Although there are no data available on the predictive value of TS mRNA and protein expressions in NSCLC response to pemetrexed, it has been hypothesized that the higher expression of TS observed in SCCs explains the lower rate of response to pemetrexed in this NSCLC type (4).

When we correlated the two potential markers, FR α and TS protein expression, with response to pemetrexed, we found that in SCCs the expression of nuclear TS had a significant inverse correlation with expression of membrane FR α , and most TS-positive SCCs (79%) lacked membrane FR α immunostaining. Furthermore, we speculate that the more frequent

occurrence of the FR α membrane-negative/TS nuclear-positive expression pattern in lung SCCs than in adenocarcinomas could be associated with the lower response rate to pemetrexed in this tumor type. In contrast, in adenocarcinomas the most frequent pattern detected was FR α membrane-positive/TS nuclear-negative which could correlate with the higher response to this drug in gystology type.

In summary, our findings indicate that membrane transporter FR α and RFC1 proteins are frequently overexpressed in NSCLC tissues. The higher level of FR α in adenocarcinomas than in SCCs may help explain differences in efficacy of antifolate chemotherapy between these tumor types. We postulate that this information may be useful in selecting which patients with NSCLC may benefit from and should receive treatment with antifolate inhibiting agents, including pemetrexed.

REFERENCES

1. Pfister DG, Johnson DH, Azzoli CG, *et al.* American Society of Clinical Oncology treatment of unresectable non-small-cell lung cancer guideline: update 2003. *J Clin Oncol* 2004;22: 330-53.
2. Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon JP, Vansteenkiste J. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 2004;350: 351-60.
3. Huang Y. Pharmacogenetics/genomics of membrane transporters in cancer chemotherapy. *Cancer Metastasis Rev* 2007;26: 183-201.
4. Scagliotti GV, Parikh P, von Pawel J, *et al.* Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naïve patients with advanced-stage non-small-cell lung cancer. *J Clin Oncol* 2008;26: 3543-51.
5. Scagliotti GV, Ceppi P, Capelletto E, Novello S. Updated clinical information on multitargeted antifolates in lung cancer. *Clin Lung Cancer* 2009;10 Suppl 1: S35-40.
6. Taylor EC, Kuhnt D, Shih C, *et al.* A dideazatetrahydrofolate analogue lacking a chiral center at C-6, N-[4-[2-(2-amino-3,4-dihydro-4-oxo-7H-pyrrolo[2,3-d]pyrimidin-5-yl)ethyl]benzoyl]-L-glutamic acid, is an inhibitor of thymidylate synthase. *J Med Chem* 1992;35: 4450-4.
7. Schultz RM, Patel VF, Worzalla JF, Shih C. Role of thymidylate synthase in the antitumor activity of the multitargeted antifolate, LY231514. *Anticancer Res* 1999;19: 437-43.
8. Shih C, Habeck LL, Mendelsohn LG, Chen VJ, Schultz RM. Multiple folate enzyme inhibition: mechanism of a novel pyrrolopyrimidine-based antifolate LY231514 (MTA). *Adv Enzyme Regul* 1998;38: 135-52.
9. Weitman SD, Weinberg AG, Coney LR, Zurawski VR, Jennings DS, Kamen BA. Cellular localization of the folate receptor: potential role in drug toxicity and folate homeostasis. *Cancer Res* 1992;52: 6708-11.

10. Zhao R, Goldman ID. The molecular identity and characterization of a Proton-coupled Folate Transporter--PCFT; biological ramifications and impact on the activity of pemetrexed. *Cancer Metastasis Rev* 2007;26: 129-39.
11. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. *Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart*. Lyon: International Agency for Research on Cancer (IARC); 2004. p. 9-124.
12. Shia J, Klimstra DS, Nitzkowski JR, *et al.* Immunohistochemical expression of folate receptor alpha in colorectal carcinoma: patterns and biological significance. *Hum Pathol* 2008;39: 498-505.
13. Shigematsu H, Lin L, Takahashi T, *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005;97: 339-46.
14. Tang X, Shigematsu H, Bekele BN, *et al.* EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res* 2005;65: 7568-72.
15. Rhodes DR, Yu J, Shanker K, *et al.* ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;6: 1-6.
16. Tomida S, Koshikawa K, Yatabe Y, *et al.* Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. *Oncogene* 2004;23: 5360-70.
17. Bild AH, Yao G, Chang JT, *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439: 353-7.
18. Kim J. Prediction of Recurrence-Free Survival in Postoperative NSCLC Patients—a Useful Prospective Clinical Practice. GEO (Gene Expression Omnibus) 2007 [cited August 29, 2007]; Available from: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8894>
19. Chen HY, Yu SL, Chen CH, *et al.* A five-gene signature and clinical outcome in non-

small-cell lung cancer. *N Engl J Med* 2007;356: 11-20.

20. Chitale D, Gong Y, Taylor BS, *et al.* An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. *Oncogene* 2009;28: 2773-83.
21. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of Gene Expression Data Using BRB-Array Tools. *Cancer Inform* 2007;3: 11-7.
22. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31: e15.
23. Kelland L. The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer* 2007;7: 573-84.
24. Ishida S, Lee J, Thiele DJ, Herskowitz I. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc Natl Acad Sci U S A* 2002;99: 14298-302.
25. Song IS, Savaraj N, Siddik ZH, *et al.* Role of human copper transporter Ctr1 in the transport of platinum-based antitumor agents in cisplatin-sensitive and cisplatin-resistant cells. *Mol Cancer Ther* 2004;3: 1543-9.
26. Iwakiri S, Sonobe M, Nagai S, Hirata T, Wada H, Miyahara R. Expression status of folate receptor alpha is significantly correlated with prognosis in non-small-cell lung cancers. *Ann Surg Oncol* 2008;15: 889-99.
27. Bueno R, Appasani K, Mercer H, Lester S, Sugarbaker D. The alpha folate receptor is highly activated in malignant pleural mesothelioma. *J Thorac Cardiovasc Surg* 2001;121: 225-33.
28. Kalli KR, Oberg AL, Keeney GL, *et al.* Folate receptor alpha as a tumor target in epithelial ovarian cancer. *Gynecol Oncol* 2008;108: 619-26.
29. Hartmann LC, Keeney GL, Lingle WL, *et al.* Folate receptor overexpression is associated with poor outcome in breast cancer. *Int J Cancer* 2007;121: 938-42.
30. Jin M, Kawakami K, Fukui Y, *et al.* Different histological types of non-small cell lung

cancer have distinct folate and DNA methylation levels. *Cancer Sci* 2009.

31. Reddy JA, Allagadda VM, Leamon CP. Targeting therapeutic and imaging agents to folate receptor positive tumors. *Curr Pharm Biotechnol* 2005;6: 131-50.
32. Ebel W, Routhier EL, Foley B, *et al.* Preclinical evaluation of MORAb-003, a humanized monoclonal antibody antagonizing folate receptor-alpha. *Cancer Immun* 2007;7: 6.
33. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. *Nat Rev Cancer* 2007;7: 778-90.
34. Eberhard DA, Johnson BE, Amler LC, *et al.* Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol* 2005;23: 5900-9.
35. Peebles KA, Lee JM, Mao JT, *et al.* Inflammation and lung carcinogenesis: applying findings in prevention and treatment. *Expert Rev Anticancer Ther* 2007;7: 1405-21.
36. Janne PA, Gray N, Settleman J. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nat Rev Drug Discov* 2009;8: 709-23.
37. Adjei AA. Pemetrexed (ALIMTA), a novel multitargeted antineoplastic agent. *Clin Cancer Res* 2004;10: 4276s-80s.
38. Ceppi P, Volante M, Saviozzi S, *et al.* Squamous cell carcinoma of the lung compared with other histotypes shows higher messenger RNA and protein levels for thymidylate synthase. *Cancer* 2006;107: 1589-96.
39. Zheng Z, Li X, Schell MJ, *et al.* Thymidylate synthase in situ protein expression and survival in stage I nonsmall-cell lung cancer. *Cancer* 2008;112: 2765-73.
40. Hanauske AR, Eismann U, Oberschmidt O, *et al.* In vitro chemosensitivity of freshly explanted tumor cells to pemetrexed is correlated with target gene expression. *Invest New Drugs* 2007;25: 417-23.

FIGURE LEGENDS

Figure 1. Photomicrographs showing immunohistochemical expression of FR α , FRC1 and TS in NSCLC tissue specimens by histologic type. FR α : *A*, strong cytoplasmic and membrane expression in tumor cells; *B* and *C*, moderate expression in tumor cells; *D*, lack of expression in malignant cells. RFC1: *E* and *G*, strong cytoplasmic expression in malignant cells. TS: *F* and *H*, negative and moderate cytoplasmic and nuclear expression in tumor cells, respectively. Original magnification, $\times 200$.

Figure 2. FR α expression scores by tumor histology. In the box-plots, *black bar* indicates median scores.

Figure 3. *FOLR1* mRNA expression scores in lung adenocarcinoma by *EGFR* mutation status using two publicly microarray datasets available (20). In the box-plots, *small open box* indicates median mRNA expression scores.

Table 1. Summary of clinicopathologic features of patients with NSCLC examined for membrane transporter expression.

Feature	NSCLC Histologic Type		
	Squamous Cell Carcinoma (n = 118)	Adenocarcinoma (n = 202)	Total (n = 320)
Mean age, years (SD), (range)	68.4 (9.20), (43-90)	64.9 (11.5), (33-88)	66.2 (10.85), (33-90)
Sex			
Male	73	77	150
Female	45	125	170
Smoking status [†]			
Never	4	52	56
Ever	113	150	263
TNM stage			
I	62	134	196
II	36	25	61
III	18	36	54
IV	2	7	9

[†] Smoking status and history were not available for one patient with squamous cell carcinoma.

Table 2. Frequency of membrane transporters and thymidylate synthase (TS) immunohistochemical expression in NSCLC by tumor histology

Marker	Any Expression (Score > 0)			Average Score		
	Squamous Cell Carcinoma Positive/Total (%)	Adenocarcinoma Positive/Total (%)	<i>P</i> Value*	Squamous Cell Carcinoma Score (SD)	Adenocarcinoma	<i>P</i> Value†
FR α						
Cytoplasm	63 / 110 (57%)	152 / 174 (87%)	< 0.001	35.9 (40.3)	91.6 (66.4)	< 0.001
Membrane	29 / 110 (26%)	107 / 174 (61%)	< 0.001	11.29 (28.8)	72.2 (89.0)	< 0.001
RFC1						
Cytoplasm	110 / 112 (98%)	181 / 182 (99%)	0.56	153.2 (72.0)	162.7 (83.2)	0.34
Membrane	103 / 112 (92%)	164 / 182 (90%)	0.68	119.2 (86.1)	128.1 (95.9)	0.59
TS						
Cytoplasm	82 / 102 (80%)	130 / 165 (79%)	0.75	55.6 (42.0)	52.2 (40.1)	0.565
Nuclear	59 / 102 (58%)	58 / 165 (35%)	0.0003	13.8 (27.7)	9.3 (21.1)	0.0043

*Fisher's exact test

†Wilcoxon rank-sum test

Table 3. Frequency of membrane FR α and thymidylate synthase (TS) immunohistochemical expression in NSCLC tumor tissues by histology

Expression Pattern	Squamous Cell Carcinoma Positive (%) n = 99	Adenocarcinoma Positive (%) n = 155
FR α / TS Nuclear		
Positive / Positive	14 (14%)	29 (9%)
Positive / Negative	12 (12%)	68 (44%)
Negative / Positive	43 (43%)	25 (16%)
Negative / Negative	30 (30%)	33 (21%)
FR α / TS Cytoplasm		
Positive / Positive	17 (17%)	73 (47%)
Positive / Negative	9 (9%)	24 (16%)
Negative / Positive	62 (63%)	49 (32%)
Negative / Negative	11 (11%)	9 (6%)

Fig. 1

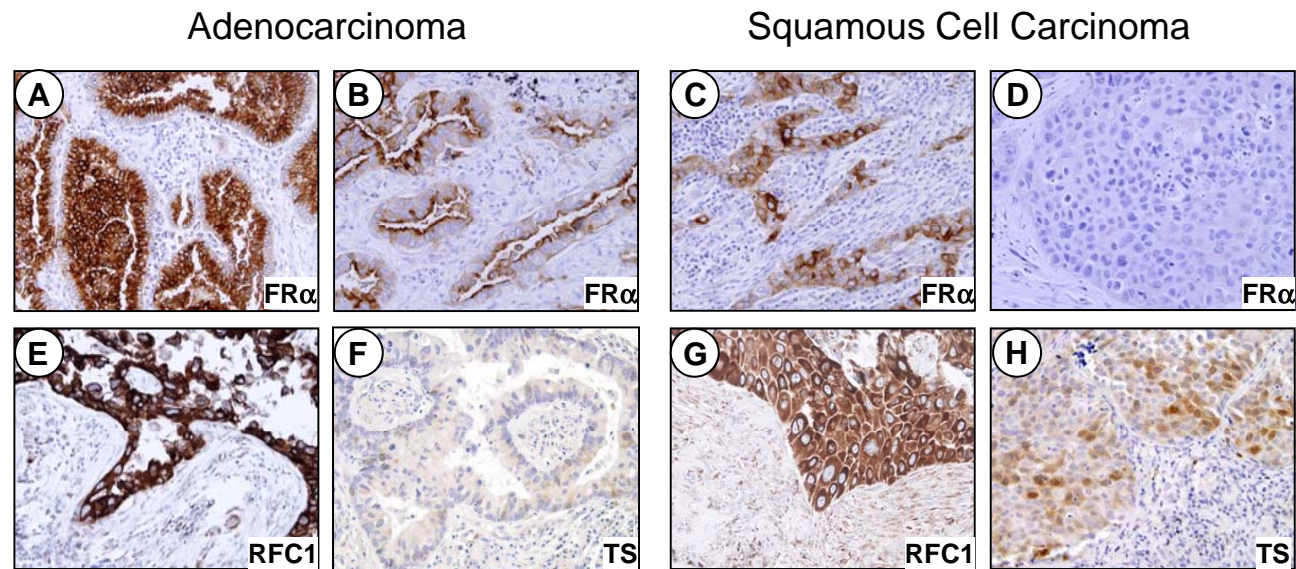


Fig. 2

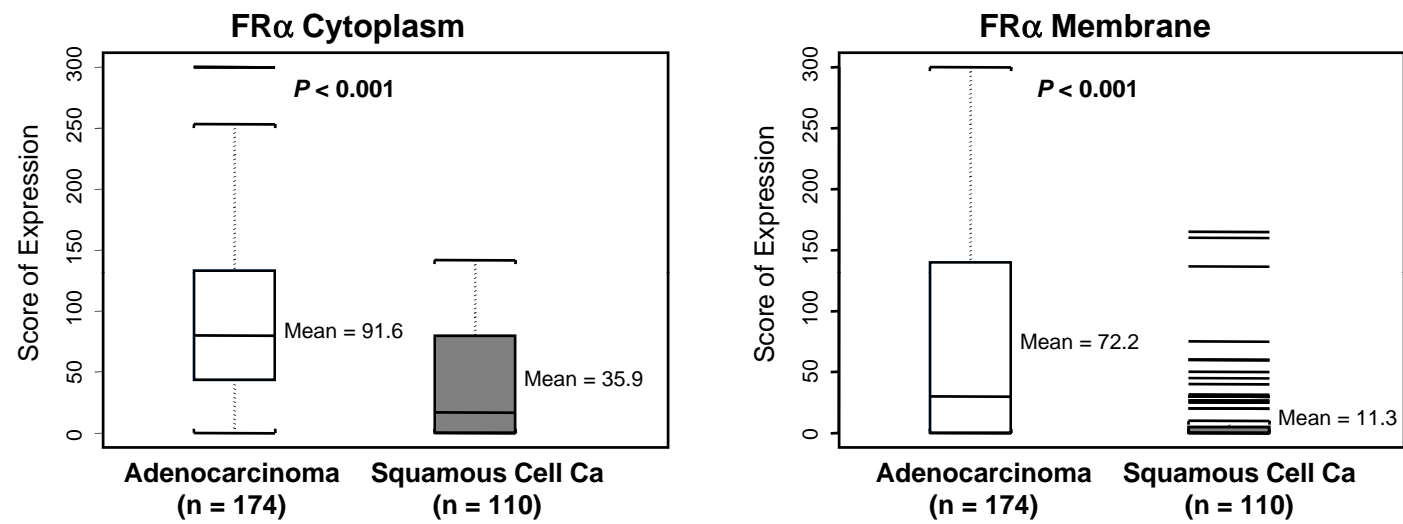
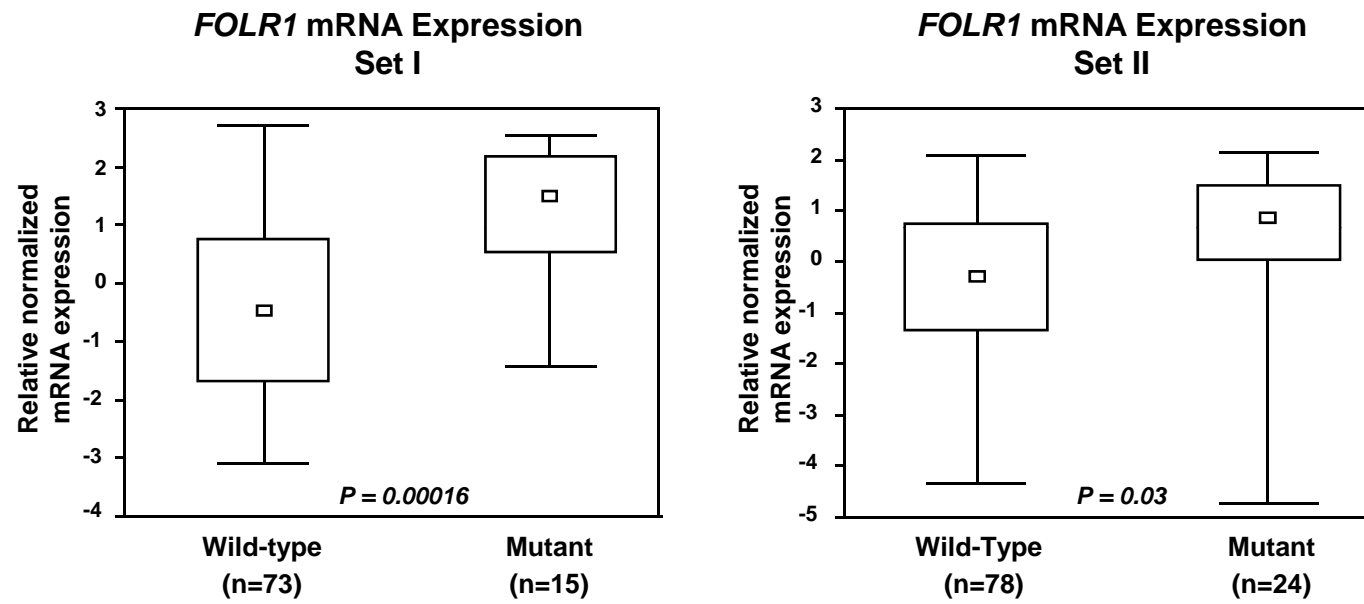


Fig. 3



Robust Gene Expression Signature from Formalin-Fixed Paraffin-Embedded Samples Predicts Prognosis of Non-Small-Cell Lung Cancer Patients

Yang Xie^{1,2}, Guanghua Xiao¹, Kevin R. Coombes⁶, Carmen Behrens⁷, Luisa Solis⁸, Gabriela Raso⁸, Luc Girard^{2,5}, Heidi Erickson⁷, Xiaosong Pan⁵, Jack Roth⁹, Kathy Danenberg¹⁰, John D. Minna^{2,3,4,5}, Ignacio I. Wistuba^{7,8}.

Authors' Affiliations: Departments of ¹Clinical Sciences, ²Simmons Cancer Center, ³Internal Medicine, ⁴Pharmacology, and ⁵Hamon Center for Therapeutic Oncology Research The University of Texas Southwestern Medical Center, Dallas, Texas; Departments of ⁶Bioinformatics and Computational Biology, ⁷Thoracic/Head and Neck Medical Oncology, ⁸Pathology and ⁹Thoracic and Cardiovascular Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas; and ¹⁰Response Genetics, Inc, Los Angeles, California.

Grant support: This study was supported in part by grants from the Department of Defense (W81XWH-07-1-0306 03 to W.K.H., J.D.M., and I.I.W.), the Specialized Program of Research Excellence in Lung Cancer Grant (P50CA70907 to J.D.M., J. R., and I.I.W.), the National Cancer Institute (Cancer Center Support Grant CA-16672), National Institute of Health (NIH UL1 RR024982 Y.X and G.X).

Running title:

Robust gene signature from fixed tissue predicts lung cancer prognosis

Abbreviations:

Non-small-cell-lung-cancer: NSCLC

Formalin-fixed paraffin-embedded: FFPE

Robust gene set: RGS

Key words: Lung Cancer Prognosis, Gene Expression Signature, Formalin Fixed Paraffin Embedded Samples

Requests for reprints: Ignacio I. Wistuba, Departments of Pathology and Thoracic/Head and Neck Medical Oncology, Unit 85, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-563-9184; Fax: 713-792-0309; E-mail: iiwistuba@mdanderson.org.

Abstract

Purpose

The requirement of frozen tissues for microarray experiments limits the clinical usage of genome-wide expression profiling using microarray technology. The goal of this study is to test the feasibility of developing lung cancer prognosis gene signatures using genome-wide expression profiling of formalin-fixed paraffin-embedded (FFPE) samples, which are widely available and provide a valuable rich source for studying the association of molecular changes in cancer and associated clinical outcomes.

Patients and Methods

We randomly selected 100 Non-Small-Cell lung cancer (NSCLC) FFPE samples with annotated clinical information from the UT-Lung SPORE Tissue Bank. We microdissected tumor area from FFPE specimen, and used Affymetrix U133 plus 2.0 arrays to attain gene expression data. After strict quality control and analysis procedures, a supervised principal component analysis was used to develop a robust prognosis signature for NSCLC. Three independent published microarray data sets were used to validate the prognosis model.

Results

This study demonstrated that the robust gene signature derived from genome-wide expression profiling of FFPE samples is strongly associated with lung cancer clinical outcomes, can be used to refine the prognosis for stage I lung cancer patients and the prognostic signature is independent of clinical variables. This signature was validated in several independent studies and was refined to 59-gene lung cancer prognosis signature.

Conclusion

We conclude that genome-wide profiling of FFPE lung cancer samples can identify a set of genes whose expression level provides prognostic information across different platforms and studies, which will allow its application in clinical settings.

Introduction

Lung Cancer is the leading cause of death from cancer for both men and women in the United States and in most parts of the world, with 5-year survival rate of 15% ¹. Non-small-cell lung cancer (NSCLC) is the most common cause of lung cancer death, accounting for up to 85% of such deaths ². Clinical-pathologic staging is the standard prognosis factor for lung cancer used in clinical practice, but does not capture the complexity of the disease so that heterogeneous clinical outcomes within the same stage are commonly seen. Several randomized clinical trials showed that adjuvant chemotherapy improves survival in resected NSCLC ³⁻⁸. The effect of adjuvant chemotherapy on prolonging survival is modest - only 4-15% improvement in 5-year survival, while such treatment is associated with serious adverse effects ^{3, 7, 9-10}. Therefore, it is of considerable clinical importance to have a robust and accurate prognostic signature for lung cancer, especially in early stage lung cancer to improve the current clinical decisions on whether an individual lung cancer patient should receive adjuvant chemotherapy or not.

Genome-wide expression profiles have been used to identify gene signatures to classify lung cancer patients with different survival outcomes ¹¹⁻²⁰. However, the requirement of frozen tissues for microarray experiments limits the clinical usage of these gene signatures. Furthermore, prognostic gene signatures for NSCLC developed by different groups show minimal overlap, and are often difficult to reproduce by independent groups ²¹⁻²². To address the problem of requirement for frozen issues, we designed this study to test the feasibility of developing lung cancer prognosis gene signatures using genome-wide expression profiling of formalin-fixed paraffin-

embedded (FFPE) samples, which are widely available and provide a valuable rich source for studying the association of molecular changes in cancer and associated clinical outcomes. We derived a prognosis signature for NSCLC from FFPE samples and validated it in several independent studies. To facilitate other researchers to reproduce all results in this study, we have provided a literate programming R package.

PATIENTS AND METHODS

Patients

The overall study design and the flow chart of the derivation and validation of the robust gene signature are described in **Figure 1**. We randomly selected 100 NSCLC FFPE samples with annotated clinical information from the UT-Lung SPORE Tissue Bank from 2001-2005. From these samples, 75 samples passed the mRNA quality control criteria. Among these 75 samples, 48 samples are adenocarcinomas and 27 are squamous cell carcinomas. The median follow-up time is 2.8 years and the maximum follow-up time is 6.9 years; the characteristics of these patients are summarized in **Supplementary Table 1**. The samples were obtained under approval of the institutional review boards at M.D. Anderson Cancer center.

Sample microdissection and RNA extraction

FFPE tumor specimens were cut into serial sections with a thickness of 10 μ m. For the pathological diagnosis, one slide was stained with H&E and evaluated by a pathologist. Other sections were stained with nuclear fast red (NFR, American MasterTech Scientific Inc., Lodi, CA) to enable visualization of histology. Tumor tissue was isolated using manual macro-dissection when the tumor area was $> 0.5 \times 0.5$ mm or laser capture microdissection (P.A.L.M. Microlaser Technologies AG, Munich,

Germany) in cases of smaller tumor areas. At least 50 mm² of tumor tissue was collected from each FFPE block. The extraction of RNA from tissue samples was done by a proprietary procedure of Response Genetics, Inc. (United States Patent Application 20090092979) designed to optimize the yield of higher molecular weight RNA fragments from FFPE specimens.

Microarray data preprocessing and quality control.

Total RNA was processed for analysis on the Affymetrix U133 plus 2.0 arrays according to Affymetrix protocols for first- and second-strand synthesis, biotin labeling and fragmentation. The quality control procedure for microarray data analysis was based on the percentage of present calls calculated by the MAS5 package. We selected arrays with at least 15% of probe sets present; 55 out of 75 arrays passed this quality control criterion and will be used for the analysis. We selected probe sets that are present on all 55 arrays; 1400 genes past this criterion. These 1400 genes were referred as the robust gene set (RGS) since the mRNA expression of these genes are robust to FFPE processing. The 55 samples and the 1400 genes were used to develop gene signatures.

After microarray analysis QC, we used the RMA background correction algorithm²³ to remove non-specific background noise. A robust regression model²⁴ was fitted to the probe level data, and the fitted expression values for the probes at the 3' end were used to summarize the probe set expression values. Quantile-quantile normalization was used to normalize all the arrays. Consortium microarray raw data¹⁵ was downloaded from the National Cancer Institute's caArray database and preprocessed by RMA background correction and quantile-quantile normalization. All gene-expression values were log-transformed (on a base 2 scale).

Supervised classification using supervised principal component analysis.

Classification was performed using supervised principal component analysis²⁵⁻²⁶, a widely used classification method in biomedical research²⁷⁻³⁰. As a supervised classification method, each prediction model was trained in a training data and then the performance was tested in an independent test dataset. We used an R package (version 2.81), Superpc (version 1.05), to implement the prediction algorithm, and the default parameters were used. The implementation details can be found in the **Supplementary Sweave Report**. The training and testing sets for each prediction model are summarized in **Supplementary Table 2**.

Survival analysis.

Overall survival time was calculated from the date of surgery until death or the last follow-up contact. Survival curves were estimated using the product-limit method of Kaplan-Meier³¹ and were compared using the log-rank test. The maximum follow-up time for the FFPE patient cohort is less than 7 years, while some patients in the consortium cohort have been followed for up to 17 years. To avoid the extrapolation of the prediction model, the comparison of survival time between predicted groups are truncated at 7 years. The analysis results without truncation can be seen in **Supplementary Sweave Report**. Univariate and multivariate Cox proportional-hazards analysis³² were also performed, with survival as the dependent variable.

Results

The robust gene set defines two tumor groups

The expression of these 1400 genes divided the 55 patients into two groups based on unsupervised clustering analysis (with euclidean distance and complete

linkage for the hierarchical clustering algorithm) (**Figure 2**). Interestingly, group 1 has significantly shorter survival time compared to group 2 (**Figure 2b**, $P=0.017$) and multivariate Cox proportional-hazards analysis showed that the association between RGS groups and survival ($P=0.012$) is independent of stage. Notably, group 1 was dominated by squamous cell carcinoma (23/28), whereas group 2 was dominated by adenocarcinomas (25/27) ($P<0.0001$) (**Supplementary Table 3**). The other clinical characteristics including gender, age and smoking status were not significantly different between the two groups. To explore whether the association between RGS groups and survival is due to the histology difference between two groups, we drew Kaplan-Meier curves by both histology and RGS groups (**Supplementary Figure 1**) and it shows clearly that RGS can distinguish high and low risk groups within both adenocarcinoma and squamous groups indicating the association of RGS groups and survival is independent of histology groups.

We used gene set enrichment analysis (GSEA) to identify the enriched gene sets in both RGS groups. Interestingly, an estrogen receptor (ER) negative signature in breast cancer³³ is enriched in RGS group 1, meanwhile, an ER positive signature in breast cancer³³ is enriched in RGS group 2 (**Figures 2c and 2d**) indicating the relationship between the ER signatures and the RGS groups. The other enriched gene sets are summarized in **Supplementary Table 4**; notably, genes enriched in group 1 are also enriched in mouse neural stem cells and embryonic stem cells.

Construct and validate RGS prognosis signatures

FFPE samples training to testing. The strong associations between RGS groups and survival outcomes motivated us to explore whether RGS expression profile can be

used to construct prognosis signature. We randomly divided 55 patients into training (25 samples) and testing (30 samples) sets, and constructed a prediction model using 1400 robust gene expression values in the training set through a supervised principle component approach ²⁵. **Figure 3a** shows that the predicted low risk group has significant longer survival time than the predicted high risk group ($P=0.013$) in the testing set. To test if this association was not random, we randomly split the data into training and testing sets 200 times, repeated the same prediction and testing procedures for each set, and found that the prognosis performance of RGS signature is significantly better than random ($P=0.02$).

Frozen samples training to testing. We then tested whether this robust gene set can be used to construct prognosis signature in frozen samples. The largest independent public available lung cancer microarray data set is the recently published NCI Director's Consortium for study of lung cancer involving 442 resected adenocarcinomas ¹⁵. From that study, Affymetrix U133A microarray data for the 1012 robust genes were excerpted with 388 less genes than our FFPE data due to the microarray platform difference. We used the same training and testing strategy as in the original analyses of these data ¹⁵ for constructing and validating prognosis signature through supervised principal component approach. The training set included samples from University of Michigan Cancer Center (UM) and Moffitt Cancer Center (HLM), and the testing set included the Memorial Sloan-Kettering Cancer Center (MSK) and Dana-Farber Cancer Institute (CAN/DF) samples. This analysis revealed that the predicted low risk group has significant longer survival time than the predicted high risk group ($P=0.000013$) in the testing dataset (**Figure 3b**).

FFPE to frozen samples and vice versa. Next, we used our FFPE and the consortium datasets as frozen samples to investigate whether the predication model built from one type of sample can be validated in another type of sample. Again, the same supervised principal component method was used to construct the prediction model. The prediction model built from FFPE samples can significantly distinguish the high and low risk groups in frozen samples (**Figure 3c**, $P=5.4 \times 10^{-7}$), and the prediction model built from frozen samples can also distinguish the high and low risk groups in FFPE samples but with marginal significance (**Figure 3d**, $P=0.068$). We also tested the performance of FFPE prediction model on four individual datasets in consortium study and found that the predicted low risk groups have longer survival time compared to the predicted high risk groups for all sets: MSKCC dataset (median survival time 6.5 vs. 3.3 years; $P=0.0093$), DFCI dataset (median survival time 5.9 vs. 0.9 years; $P=0.0076$), HLM dataset (median survival time 3.4 vs. 2.2 years; $P=0.4$) and MI data set (median survival time 5.4 vs. 2.2 years; $P=0.0011$) (**Supplementary Figure 2**).

The RGS prognosis signature is independent of clinical variables

To test whether RGS is an independent prognosis signature, we fitted a multivariate Cox regression model including RGS risk scores, age, gender, stage, smoking status, adjuvant chemotherapy usage and clinical sites as co-variables for the consortium data set. The RGS risk scores were calculated from the prediction model built from the FFPE samples set. **Table 1** shows that the RGS signature is significantly associated with the survival time after adjusting for other clinical variables ($HR=1.3$, $P=0.007$). Pathological stages based on international staging system is the most widely used and important prognosis variable for lung cancer patients³⁴, here we

tested whether RGS signature can further refine the prognosis within each stage. The RGS prognosis signature from FFPE samples was tested within each stage of the consortium dataset. The results show clearly that the RGS signature is significantly associated with survival outcome within each stage (**Figure 3e to g**; $P=0.036$ for stage I, $P=0.022$ for stage II and $P=0.021$ for stage III), indicating that the RGS signature can refine the prognosis for lung cancer patients. The RGS prognosis signature from FFPE samples was further tested for patients with or without adjuvant chemotherapy separately, and the results show clearly that the RGS signature is significantly associated with survival for both groups (**Supplementary Figure 3 a,b**; $P=0.015$ for patients with chemotherapy, $P=0.00062$ for patients without chemotherapy).

Refine to 59-gene prognosis signature

Among all the RGS genes, 131 genes are associated with survival ($P<0.05$) in the FFPE dataset, and 365 genes are associated with overall survival ($P<0.05$) in the consortium dataset by univariate Cox regression analysis. There is significant overlap between these two gene lists (**Figure 4a**; 59 common genes; $P=0.0008$, hypergeometric test). More significant genes were found in the consortium data compared to the FFPE data, which is likely due to the larger sample size ($n=442$) of the consortium dataset compared to the FFPE dataset sample size ($n=55$). Surprisingly, hazard ratios from the two datasets are very consistent with each other. All 59 genes have the same direction of effects (positive or negative) on the survival between the two data sets and the hazard ratios from two datasets are highly correlated (Pearson's correlation = 0.86) (**Figure 4b**), indicating the high consistency of expressions of these genes across datasets. These results motivated us to hypothesize that these 59 genes

(**Supplementary Table 5**) alone can be used for lung cancer prognosis. To test this hypothesis, we applied supervised principal component analysis to these 59 genes using the FFPE dataset to construct a 59-gene prognosis signature. Because the selection of these 59 genes used information from both FFPE and consortium datasets, we used another two independent lung cancer datasets including the Bild et al. (n=63)¹¹ dataset and the Bhattacharjee et al. dataset (n=117)³⁵ downloaded from the literature to validate our 59-gene signature. The 59-gene prediction model built from FFPE samples can significantly distinguish the high and low risk groups for both the Bhattacharjee et al and Bild et al. data sets (**Figure 5a**, $P=0.016$ and **Figure 5c**, $P=0.011$, respectively). Furthermore this signature can also significantly distinguish the high and low risk groups within stage I patients for both datasets (**Figure 5b 5d**) indicating this 59-gene signature can refine the prognosis for lung cancer patients within stage I patients. We also found that 59-gene prediction model built from the consortium dataset can also distinguish the high and low risk groups for the Bild et al. and Bhattacharjee et al datasets (**Supplementary Figure 4 a-d**).

To understand the potential biological relevance of these 59 genes significantly associated with survival in the FFPE and consortium data sets, we used Ingenuity Pathway Analysis (IPA) to explore which known regulatory networks are enriched in this 59-gene set. IPA analysis revealed the most significant molecular network to be cancer, tumor morphology, and respiratory disease. This network (**Figure 4c**) includes 14 genes of the 59-gene set and is centered on transcription factors *HNF4A*, *HNF1A*, and *ONECUT1* (*HNF6A*). This hepatocellular network has been implicated in hepatocellular carcinoma as determined by *in vitro* study³⁶ and molecular interactions in this network are putatively involved in lung cancer survival.

Discussion

In this study, we tested the feasibility of deriving a lung cancer prognosis gene signature from formalin-fixed paraffin-embedded tumor samples based on genome-wide mRNA expression profiling. Although RT-PCR method have been used to measure gene expression level from FFPE samples³⁷⁻³⁹, the selection of genes for testing are limited to the current knowledge base which is incomplete and inconsistent⁴⁰. Due to degradation and chemical alteration of RNA extracted from FFPE samples, the use of microarray analysis of gene expression from FFPE samples has been hampered⁴⁰. New technology and methodologies developed to extract RNA from FFPE samples coupled with new array platforms have made it possible to measure gene expression from FFPE samples^{37, 41-44}. A recent study demonstrated the feasibility of using DNA-mediated annealing, selection, extension and ligation (DASL) arrays with 6100 preselected genes to profile mRNA expression from hepatocellular carcinoma tissue⁴⁵. No prognosis signature for other types of cancer has been developed using microarray analysis of gene expression from FFPE extracted RNA. In this study, we built a robust gene signature for NSCLC based on microarray analysis of FFPE samples. We claim this is a robust gene signature because it has been validated in 6 independent published datasets including 4 sets from the consortium study and 2 additional studies from DFCI and Duke. We also built a prediction model using the same set of robust genes from frozen samples and validated the model in both frozen and FFPE samples.

Most published gene signatures identified from different studies are usually very different and with little overlap. However, we found that there is significant overlap

among the robust genes associated with survival outcomes between the FFPE dataset and the consortium dataset ($P=0.008$). More impressively, the hazard ratios, indicating the strength of the association of genes expression and survival time, are highly consistent between two independent datasets. Our interpretation for this consistency across studies is that the gene expression variation across studies is a major contribution to signature differences across studies. In this study, we used strict quality steps to exclude genes that were not expressed in our FFPE samples. This allowed for analysis of the remaining genes which had more stable expression patterns and were more robust to environment changes. Validation of our novel 59-gene signature prognostic for NSLC survival in two additional independent datasets further confirmed the robustness of these genes.

Besides the prognostic signature, the predictive signatures to determine the optimal chemotherapy regimen for individual patients also have tremendous clinical benefit. Tumor samples from clinical trials data are important to develop predictive signatures to reduce the selection bias for evaluating treatment efficacy within signature groups. However, very limited frozen tumor samples are available from completed clinical trials, so no predictive signature for chemotherapy response developed from clinical trial data has been confirmed yet. Our study demonstrated the feasibility of using FFPE samples for genome-wide mRNA profiling. Therefore, this study provides an important step to construct and validate predictive signatures for chemotherapy response using the available FFPE samples from clinical trials in the future.

FIGURE LEGENDS

Figure 1. (a) Flow chart of the derivation and validation of the robust gene signature from formalin-fixed and paraffin-embedded samples collected from M.D. Anderson UT-Lung Cancer SPORE tissue bank. **(b)** Flow chart of the derivation and validation of 59-gene prognosis signature.

Figure 2. Microarray analysis of the gene-expression profiles from formalin-fixed and paraffin-embedded (FFPE) lung tumor samples. **(a)** Unsupervised cluster analysis of the 55 FFPE lung cancer patient cohort using the expression profile of 1400 robust genes pass the microarray quality control criterion. Horizontal and vertical axes represent robust genes and lung cancer patient clusters, respectively. **(b)** Kaplan-Meier plot showing the association of the expression of robust genes with patient survival *P*-values were obtained using the log-rank test. Red color represents sample Cluster I and black color represents sample Cluster II defined by unsupervised clustering algorithm using robust gene profiling data. • indicates censored samples. Gene set enrichment analysis found that the ER negative signature derived from breast cancer patients is enriched in group 1 defined by RGS expression **(c)**, and the ER positive signature derived from breast cancer patients is enriched in group 2 defined by RGS expression **(d)**. The y axis shows running ES scores for the specific gene set on the 1400 pre-ranked genes. The x axis shows the rank in the ordered dataset. The vertical lines represent the locations of the genes that in the specific gene set.

Figure 3. Kaplan-Meier plots showing the predictive power of the robust gene signature in training and testing sets from different sets. 55 FFPE tumor samples from M. D. Anderson Cancer Center were randomly divided into training (25 samples) and testing (30 samples) sets **(a)**. Independent validation of the robust gene signature in the 442-frozen-sample cohort from multi-institute consortium. The microarray data sets were divided into two groups, one for the training and the other for the testing cohort according to the original paper **(b)**. The training data is 55 FFPE tumor samples and the testing data set is 442-frozen-sample cohort from multi-institute consortium. The testing was done for all patients **(c)**, stage I patients **(e)**, stage II patients **(f)** and stage III patients **(g)** separately. The training data is 442-frozen-sample cohort from consortium data and the testing data is 55 FFPE samples from M.D. Anderson Cancer Center **(d)**. *P* values were obtained by the log-rank test. Red and black lines represent predicted high- and low-risk groups, respectively. • indicates censored samples.

Figure 4. Comparison of individual gene effect across FFPE samples from M, D. Anderson Cancer Center and 442 frozen samples from consortium. **(a)** Venn-diagram of genes associated with overall survival ($P < 0.05$ in univariate Cox regression models). It shows 59 genes are significantly associated with survival in both FFPE data and consortium data. **(b)** The hazard ratios from univariate Cox regression models for the 59 genes common in both sets are consistent between FFPE set and consortium set. **(c)** Regulatory gene and protein interaction networks defined by the 59 predictors. Computational molecular interaction network prediction based on genes and proteins associated with the significant pathways in the Ingenuity Pathways Knowledge Base (IPKB) by Ingenuity Pathways Analysis (IPA). Interactions between the different nodes

are given as solid (direct interaction) and dashed (indirect interaction) lines (edges) with various colors for the different interaction types. This network received the highest score by IPA and is mostly centered on the transcription factors *HNF4A* and *HNF1A*, and *ONECUT1*. The genes with shades are the genes belonging to 59-gene signature.

Figure 5. Kaplan-Meier plots showing the predictive power of the 59-gene signature for two independent validation sets. The training data is 55 FFPE tumor samples from M.D. Anderson Cancer Center and the testing data set is frozen samples from lung cancer patients from Bhattacharjee et al ³⁵ dataset **(a)**, the stage I patients from Bhattacharjee et al dataset **(b)**, frozen samples from lung cancer patients from Bild et al ¹¹ dataset **(c)**, and the stage I patients from Bild et al dataset **(d)**. *P* values were obtained by the log-rank test. Red and black lines represent predicted high- and low-risk groups, respectively. • indicates censored samples.

SUPPLEMENTARY FIGURES

Supplementary Figure 1. Kaplan-Meier survival plots for different histology (a) and histology combined with RGS groups defined by expression of 1400 robust genes in M.D. Anderson Cancer Center patients (b). *P* values were obtained by log-rank test. Red and black lines represent predicted high- and the low-risk groups, respectively. Solid and dash lines represent adenocarcinoma and squamous groups, respectively.

• indicates censored samples.

Supplementary Figure 2. Kaplan-Meier plots showing the predictive power of the robust gene signature from different sets. The training data is 55 FFPE tumor samples and the testing data sets are frozen samples from 4 institutions in the consortium data: MSKCC (a), DFCI (b), MI (c) and HLM (d). *P*-values were obtained by the log-rank test. Red and black lines represent predicted high- and low-risk groups, respectively. • indicates censored samples.

Supplementary Figure 3. Kaplan-Meier plots showing the predictive power of the robust gene signature for patients with (a) and without (b) chemotherapy. The training data is 55 FFPE tumor samples and the testing data sets are frozen samples the consortium data. *P*-values were obtained by the log-rank test. Red and black lines represent predicted high- and low-risk groups, respectively. • indicates censored samples.

Supplementary Figure 4 Kaplan-Meier plots showing the predictive power of the 59-gene signature for two independent validation sets. The training data is the 442-frozen-sample cohort from multi-institute consortium dataset, and the testing data set is frozen samples from Bhattacharjee et al dataset ³⁵ **(a)**, the stage I patients from Bhattacharjee et al dataset **(b)**, frozen samples from lung cancer patients from Bild et al dataset ¹¹ **(c)**, and the stage I patients from Bild et al dataset **(d)**. *P* values were obtained by the log-rank test. Red and black lines represent predicted high- and low-risk groups, respectively. • indicates censored samples.

SUPPLEMENTARY METHODS

RNA extraction. The extraction of RNA from tissue samples was done by a proprietary procedure of Response Genetics, Inc. (United States Patent Application 20090092979) designed to optimize the yield of higher molecular weight RNA fragments from FFPE specimens. In brief, tissue samples to be extracted are incubated for 16 hr at 50 C in a solution of 20 mM Tris, pH 8.0, 3.6 mM EDTA, 1% SDS containing 1 µg/mL of proteinase K. After mixing this solution with phenol/chloroform/isoamyl alcohol (PCI) and glycogen, the RNA is precipitated by addition of isopropanol, centrifuged down and washed with ethanol. A second precipitation is then carried out to further purify the RNA from contaminating DNA. The RNA pellet is mixed with 0.5% sarcosine - guanidine isothiocyanate (GITC), 20 mM DTT solution, 50 µL of 5 mM Tris and 50 µL of 2M sodium acetate followed by PCI extraction and precipitation of the RNA with glycogen and isopropanol. After washing it with ethanol, the RNA pellet is dissolved in 5 mM Tris, pH 7.5. The concentration of nucleic acids in the preparation (the yield) is measured by reading the absorbance at 260 nm in a Nanodrop 2000 (Thermo Scientific) and the purity of the RNA is estimated by the 260/280 absorbance ratio. A quality control procedure is then done (see below) to ascertain the amount of “higher quality” (250 bp and greater) RNA and the amount of DNA contamination.

RNA quality control. RNA isolated from FFPE tissue is generally quite extensively fragmented. We found that a certain level of 300 base-length or greater RNA fragments in the RNA preparation was critical to the success of microarray analysis of FFPE tissues. Therefore, each RNA isolation was checked to determine if its content of the higher molecular weight RNA reached the required level. The isolated RNA is

converted to cDNA as previously described⁴⁶ except that oligo dT primers are used. Thus, only mRNA fragments containing a 3'-oligo A tail will be extended and converted to cDNA, thereby providing a starting point from which to measure fragment length. PCR amplification of β -actin mRNA was used to represent the total population of mRNA. Primers were chosen to amplify approximately 85 base-length segments of the β -actin gene representing locations approximately 300 and 400 bases from the 3'-end of the mRNA. The PCR primer sequences were as follows: for the segment 206-293 bases from 3' end, GCCACCCCACTTCTCTCTAAGG, ATAATTTACACGAAAGCAATGCTATCAC, 6FAM-ATGGCCCAGTCCTCTCCCAAGTCCA; for the segment 322-407 from 3' end, CATCCCCCAAAGTTCACAATGT, CAATGCATCTCATATTTGGAATGACT, 6FAM-CAACAATGTGCAATCAAAGTCCTCGGC. PCR was carried out as previously described⁴⁶. Based on our experience, we judged the RNA preparation to be suitable for the microarray if the β -actin Ct for both amplicons was 31 cycles or less. To assess DNA contamination, PCR was performed on the RNA solution without first performing reverse transcription (*i.e.*, a "no-RT" control). Thus, any signal here would be due to DNA contamination. We judged a preparation to be suitable for the microarray if the β -actin Ct in the no-RT control was at least 3 cycles less than that of the cDNA preparation.

Hybridization of RNAs to oligonucleotide arrays.

Total RNA was processed for analysis on the Affymetrix U133 plus 2.0 arrays according to Affymetrix protocols for first- and second-strand synthesis, biotin labeling and fragmentation. Genechip®Two-Cycle cDNA Synthesis kit from Affymetrix was used according to the manufacturer's protocol for amplification of signal. Samples

were hybridized to the arrays overnight at 45°C, rotating at 60 r.p.m., and were washed and scanned as per the Affymetrix protocol.

Gene set enrichment and pathway analysis. The fold changes of expression values between RGS group 1 and group 2 was used to rank the genes and then applied to Gene set enrichment analysis (GSEA) software to test which gene sets are enriched in RGS group 1 and group 2. GSEA and Ingenuity Pathway Analysis were applied to the 1400-gene list to test which gene sets and pathways are enriched in the 1400 robust gene set.

Sweave report. In order to allow others to reproduce any or all parts of our statistical analyses, we used Sweave to generate our program and report. Sweave is a literate programming R package for reproducible research; it embeds R code inside LaTeX and replaces the code with the result of running the code. The complete Sweave report is included in the supplemental material.

REFERENCES

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71-96.
2. Tsuboi M, Ohira T, Saji H, et al. The present status of postoperative adjuvant chemotherapy for completely resected non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* 2007;13:73-7.
3. Strauss GM, Herndon J, Maddaus MA, et al. Randomized Clinical Trial of adjuvant chemotherapy with paclitaxel and carboplatin following resection in Stage IB Non-Small Cell Lung Cancer (NSCLC): Report of Cancer and Leukemia Group B (CALGB) Protocol 9633. *J Clin Oncol (Meeting Abstracts)* 2004;22:7019-.
4. Douillard JY, Rosell R, De Lena M, et al. Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. *Lancet Oncol* 2006;7:719-27.
5. Kato H, Ichinose Y, Ohta M, et al. A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. *N Engl J Med* 2004;350:1713-21.
6. The International Adjuvant Lung Cancer Trial Collaborative Group. Cisplatin-Based Adjuvant Chemotherapy in Patients with Completely Resected Non-Small-Cell Lung Cancer. *N Engl J Med* 2004;350:351-60.
7. Winton T, Livingston R, Johnson D, et al. Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. *N Engl J Med* 2005;352:2589-97.
8. Strauss GM, Herndon JE, II, Maddaus MA, et al. Adjuvant Paclitaxel Plus Carboplatin Compared With Observation in Stage IB Non-Small-Cell Lung Cancer: CALGB 9633 With the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. *J Clin Oncol* 2008;26:5043-51.
9. Douillard JY. [Adjuvant chemotherapy for non-small cell lung cancer--which agents for which patients?]. *Rev Mal Respir* 2005;22:8S118-23.
10. Olaussen KA, Mountzios G, Soria JC. ERCC1 as a risk stratifier in platinum-based chemotherapy for nonsmall-cell lung cancer. *Curr Opin Pulm Med* 2007;13:284-9.
11. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439:353-7.
12. Boutros PC, Lau SK, Pintilie M, et al. Prognostic gene signatures for non-small-cell lung cancer. *Proc Natl Acad Sci U S A* 2009;106:2824-8.
13. Chen HY, Yu SL, Chen CH, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 2007;356:11-20.
14. Lu Y, Lemon W, Liu PY, et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med* 2006;3:e467.
15. Shedden K, Taylor JM, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14:822-7.
16. Sun Z, Wigle DA, Yang P. Non-overlapping and non-cell-type-specific gene expression signatures predict lung cancer survival. *J Clin Oncol* 2008;26:877-83.
17. Lau SK, Boutros PC, Pintilie M, et al. Three-gene prognostic classifier for early-stage non small-cell lung cancer. *J Clin Oncol* 2007;25:5562-9.

18. Hsu DS, Balakumaran BS, Acharya CR, et al. Pharmacogenomic strategies provide a rational approach to the treatment of cisplatin-resistant patients with advanced cancer. *J Clin Oncol* 2007;25:4350-7.
19. Hayes DN, Monti S, Parmigiani G, et al. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol* 2006;24:5079-90.
20. Bremnes RM, Veve R, Gabrielson E, et al. High-throughput tissue microarray analysis used to evaluate biology and prognostic significance of the E-cadherin pathway in non-small-cell lung cancer. *J Clin Oncol* 2002;20:2417-28.
21. Coombes KR, Wang J, Baggerly KA. Microarrays: retracing steps. *Nat Med* 2007;13:1276-7; author reply 7-8.
22. Ioannidis JPA, Allison DB, Ball CA, et al. Repeatability of published microarray gene expression analyses. *Nat Genet* 2009;41:149-55.
23. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003;19:185-93.
24. Huber PJ. 1972 Wald Lecture - Robust Statistics - Review. *Ann Math Stat* 1972;43:1041-8.
25. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004;2:E108.
26. Breiman L, Friedman J, Stone JC, Olshen RA. *Classification and Regression Trees*. Chapman & Hall/CRC 1984.
27. Garzotto M, Beer TM, Hudson RG, et al. Improved detection of prostate cancer using classification and regression tree analysis. *J Clin Oncol* 2005;23:4322-9.
28. Hess KR, Abbruzzese MC, Lenzi R, Raber MN, Abbruzzese JL. Classification and regression tree analysis of 1000 consecutive patients with unknown primary carcinoma. *Clin Cancer Res* 1999;5:3403-10.
29. Koziol JA, Zhang JY, Casiano CA, et al. Recursive partitioning as an approach to selection of immune markers for tumor diagnosis. *Clin Cancer Res* 2003;9:5120-6.
30. Valera VA, Walter BA, Yokoyama N, et al. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. *Ann Surg Oncol* 2007;14:34-40.
31. Kaplan ELM, P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 1958;53:457-81.
32. Collett D. *Modelling survival data in medical research* Chapman & Hall/CRC 2003.
33. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-6.
34. Mountain CF. The new International Staging System for Lung Cancer. *Surg Clin North Am* 1987;67:925-35.
35. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790-5.
36. Hatzis P, Talianidis I. Regulatory mechanisms controlling human hepatocyte nuclear factor 4alpha gene expression. *Mol Cell Biol* 2001;21:7320-30.

37. Farragher SM, Tanney A, Kennedy RD, Paul Harkin D. RNA expression analysis from formalin fixed paraffin embedded tissues. *Histochem Cell Biol* 2008;130:435-45.
38. Cronin M, Pho M, Dutta D, et al. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol* 2004;164:35-42.
39. Gianni L, Zambetti M, Clark K, et al. Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol* 2005;23:7265-77.
40. van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 2008;452:564-70.
41. Loudig O, Milova E, Brandwein-Gensler M, et al. Molecular restoration of archived transcriptional profiles by complementary-template reverse-transcription (CT-RT). *Nucleic Acids Res* 2007;35:e94.
42. Penland SK, Keku TO, Torrice C, et al. RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest* 2007;87:383-91.
43. Ravo M, Mutarelli M, Ferraro L, et al. Quantitative expression profiling of highly degraded RNA from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. *Lab Invest* 2008;88:430-40.
44. Roberts RA, Sabalos CM, LeBlanc ML, et al. Quantitative nuclease protection assay in paraffin-embedded tissue replicates prognostic microarray gene expression in diffuse large-B-cell lymphoma. *Lab Invest* 2007;87:979-97.
45. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 2008;359:1995-2004.
46. Schneider S, Uchida K, Salonga D, Yochim JM, Danenberg KD, Danenberg PV. Quantitative determination of p16 gene expression by RT-PCR. *Methods Mol Biol* 2004;281:91-103.

Robust mRNA Gene Signature from Formalin Fixed Paraffin Embedded Samples Predicts Prognosis of Lung Cancer Patients

Figure 1.

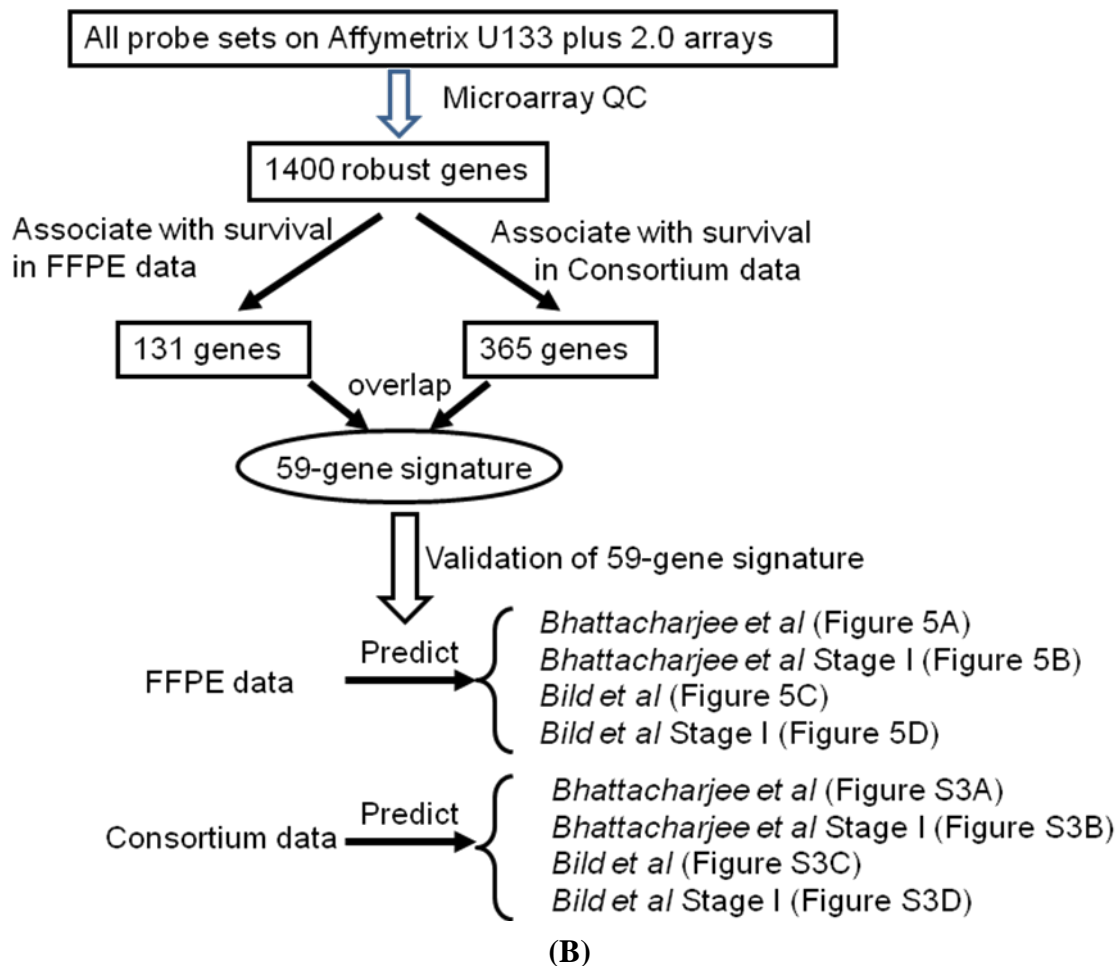
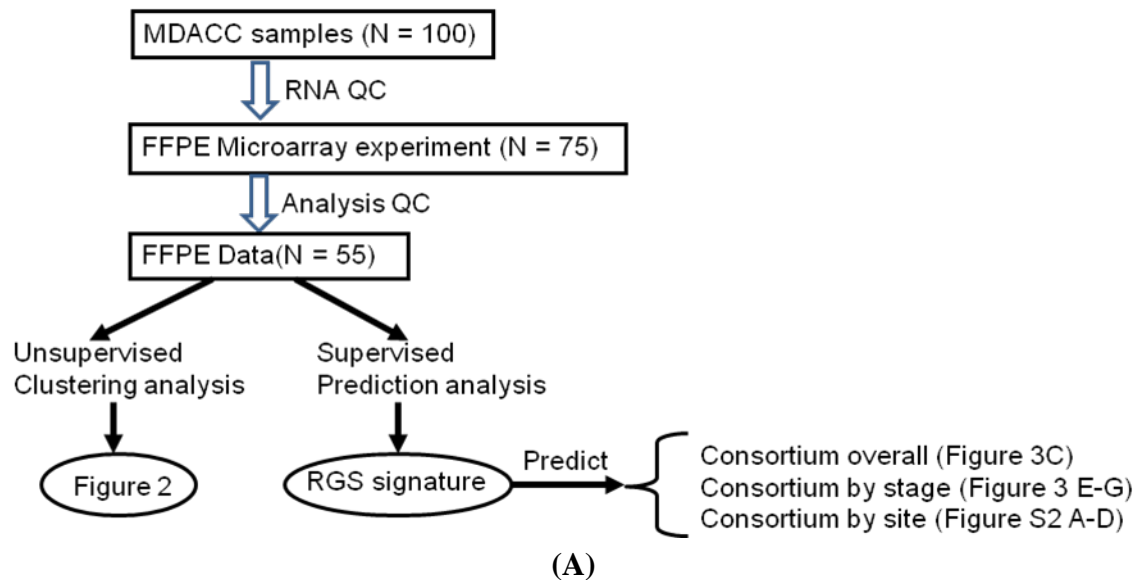
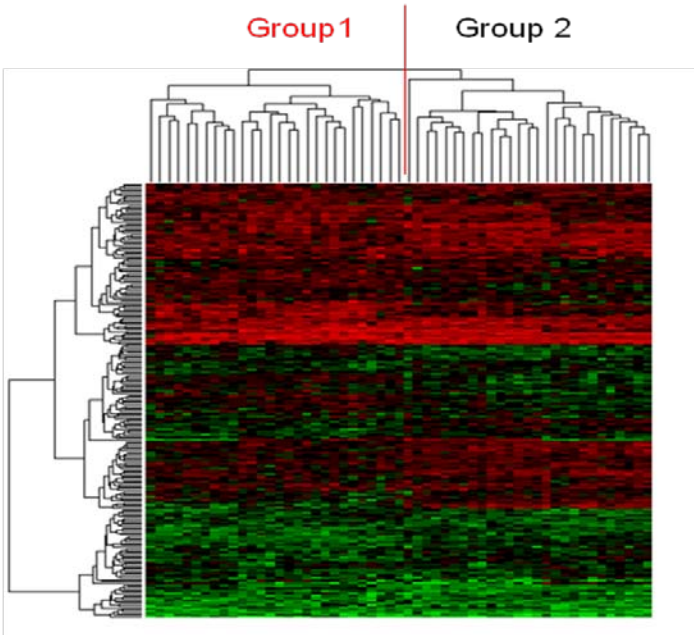
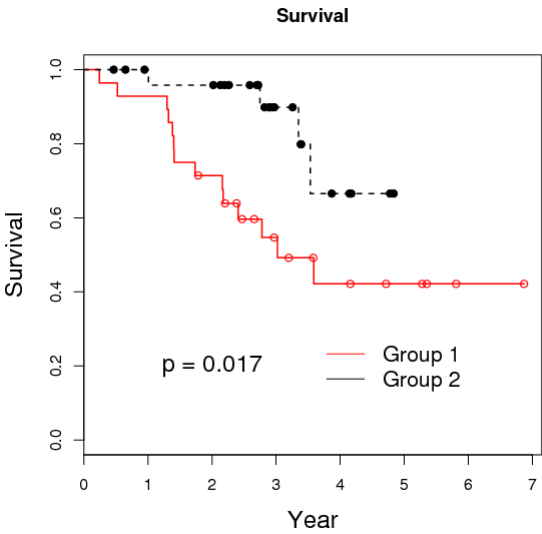


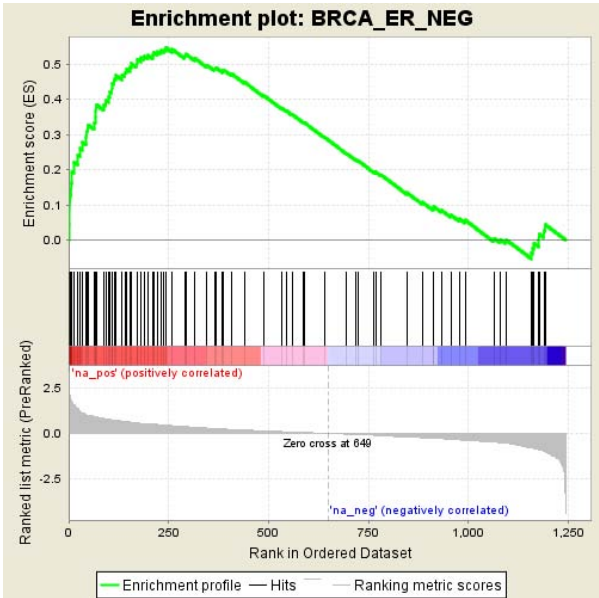
Figure 2



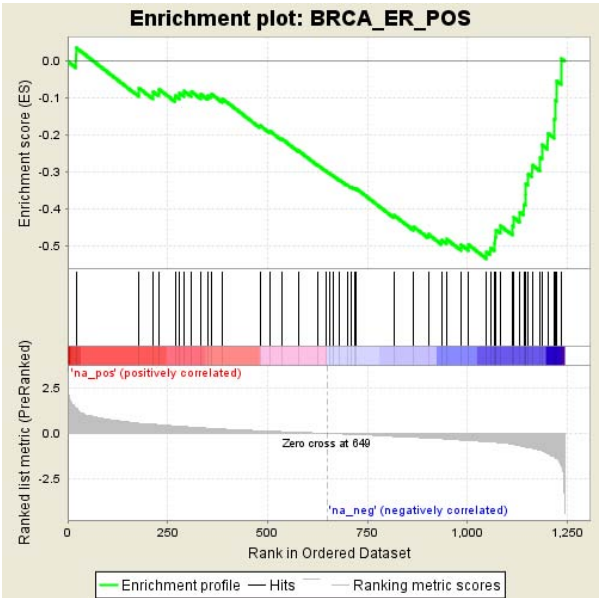
(A).



(B).

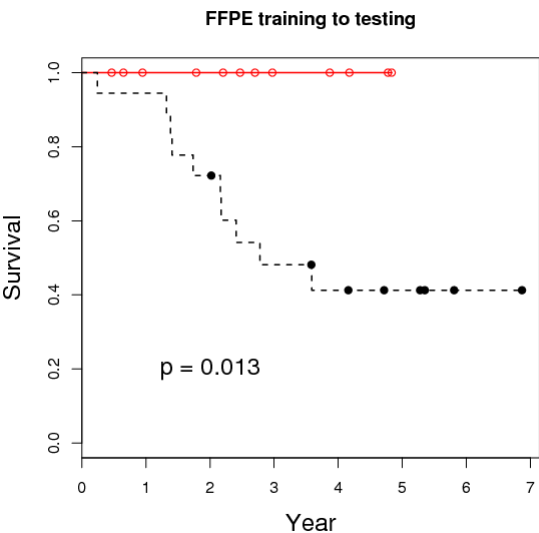


(C).

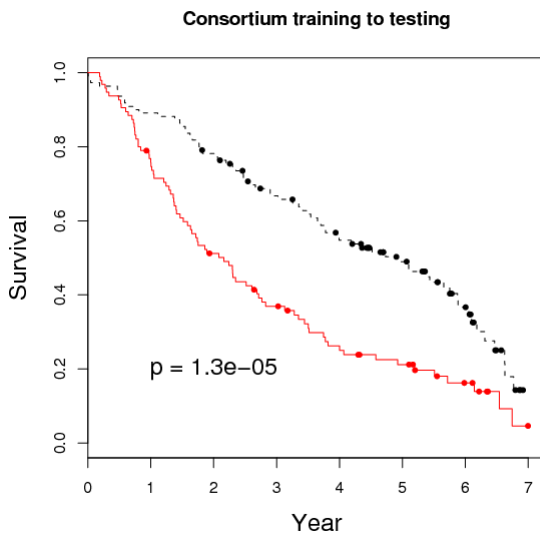


(D).

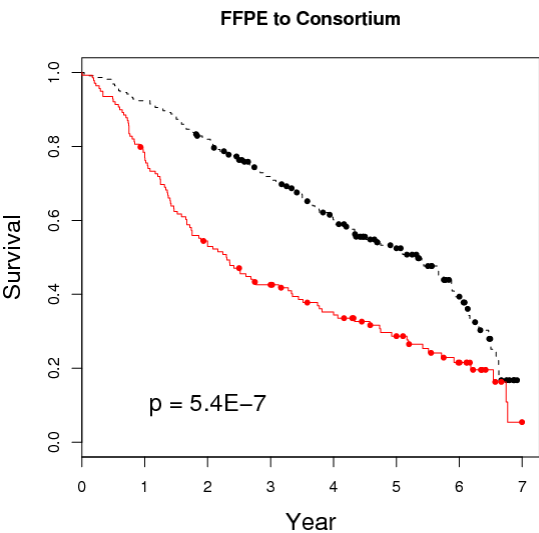
Figure 3



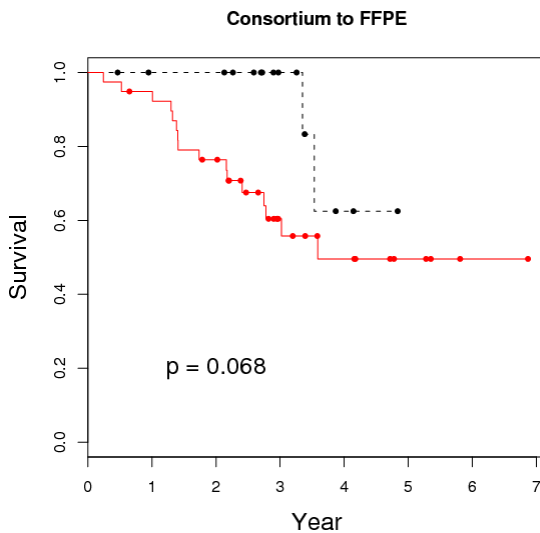
(A).



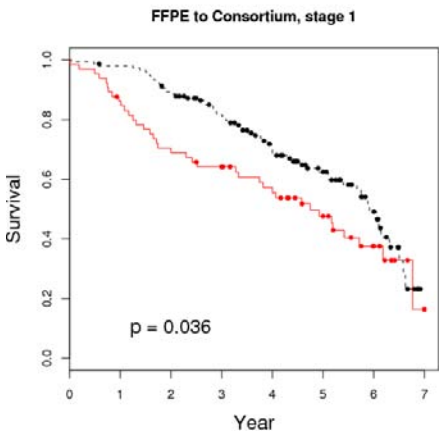
(B)



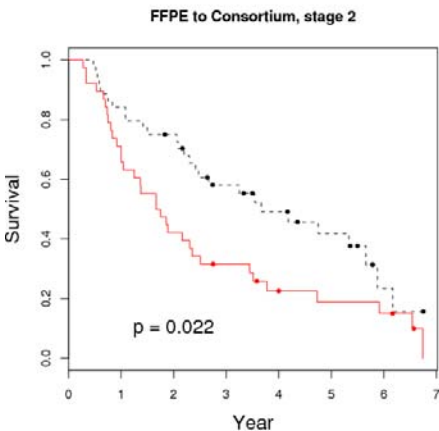
(C).



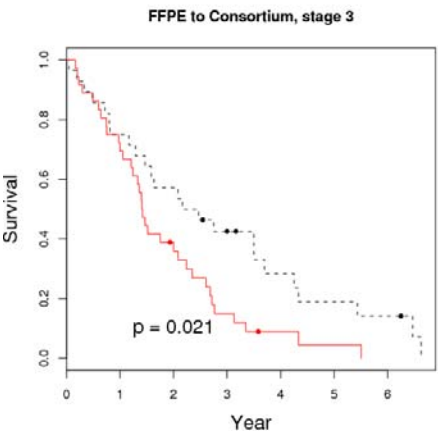
(D)



(E)



(F)



(G)

Figure 4.

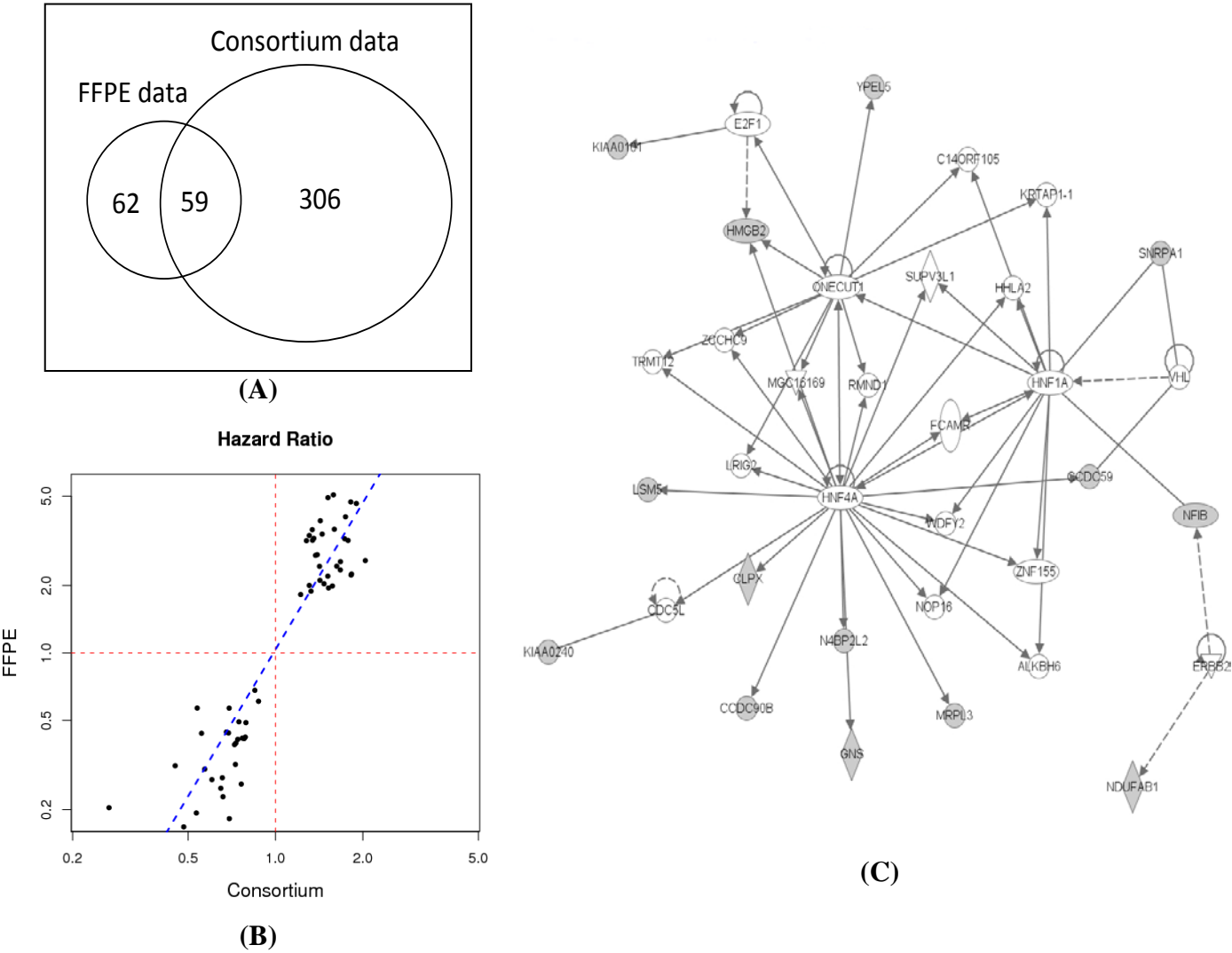
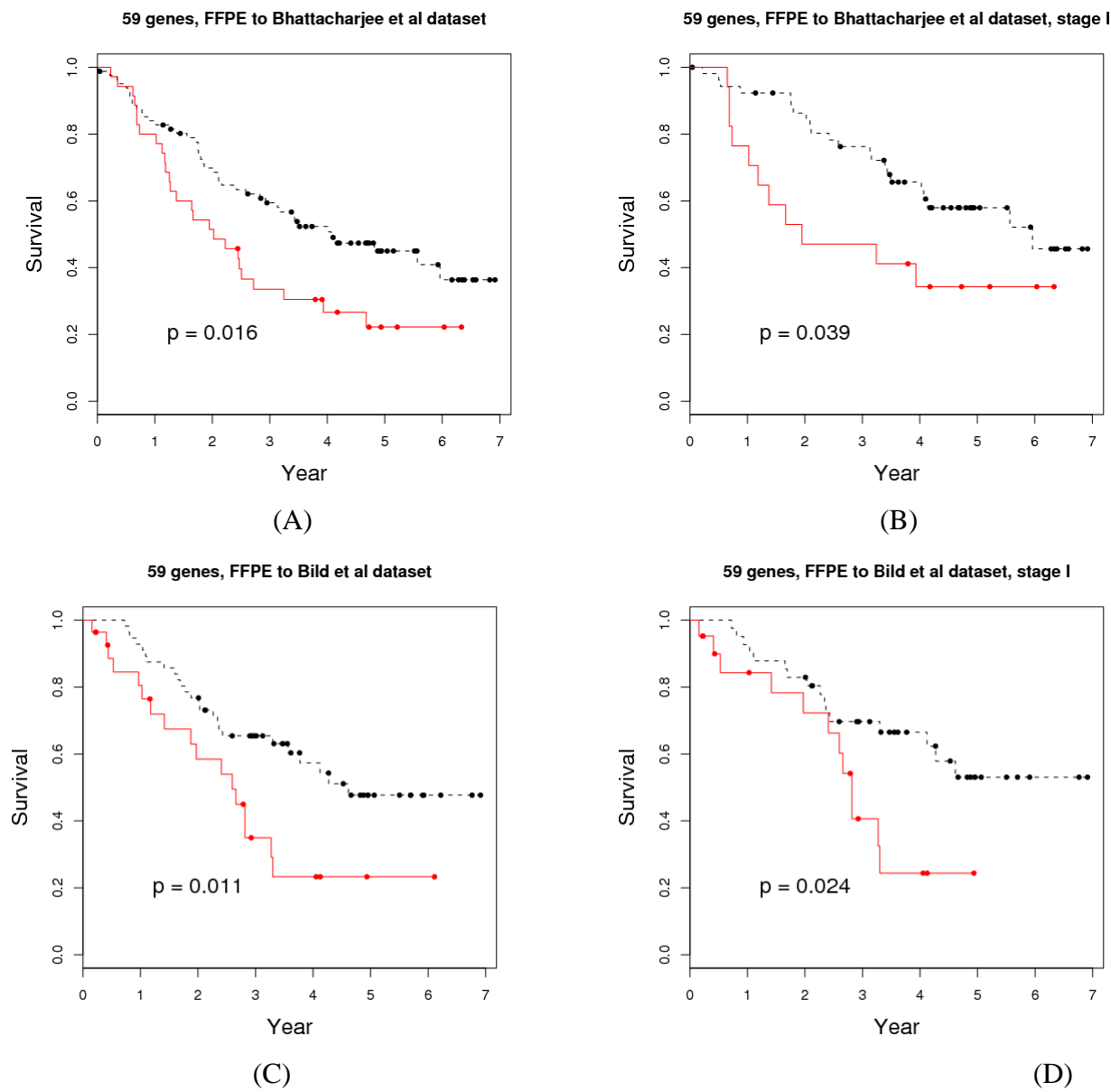
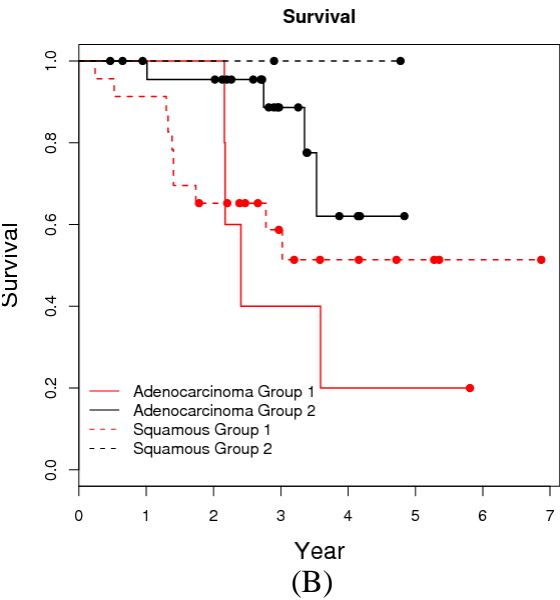
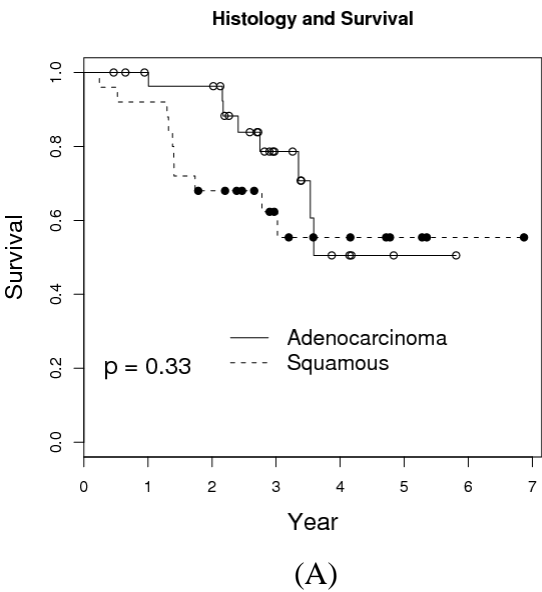


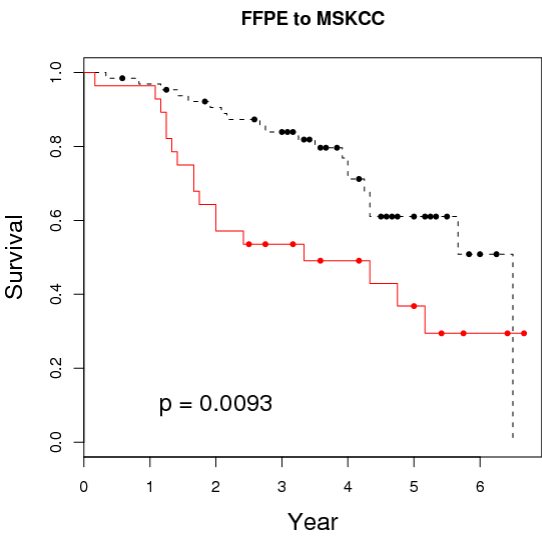
Figure 5.



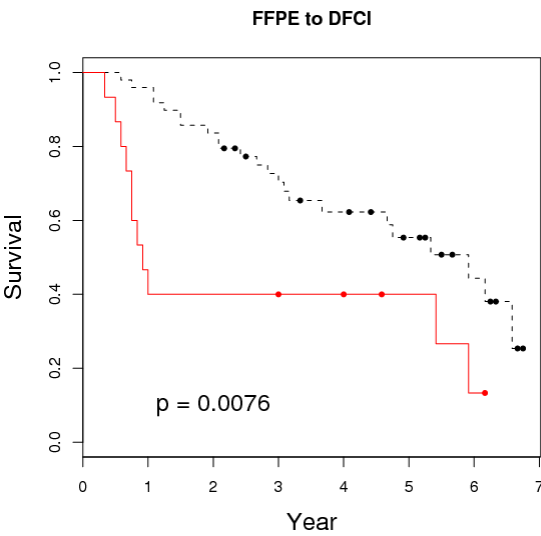
Supplementary Figure 1.



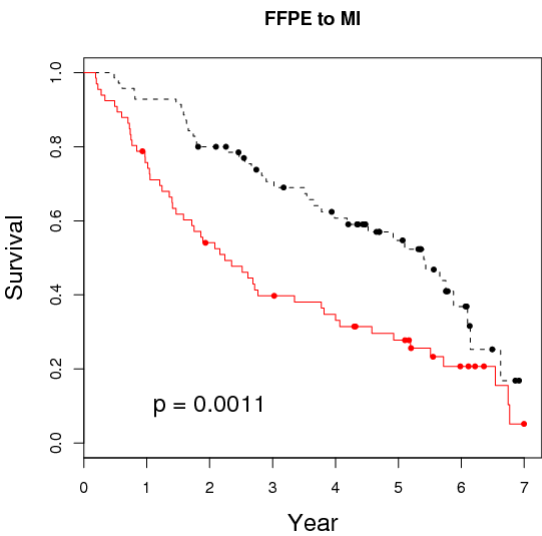
Supplementary Figure 2.



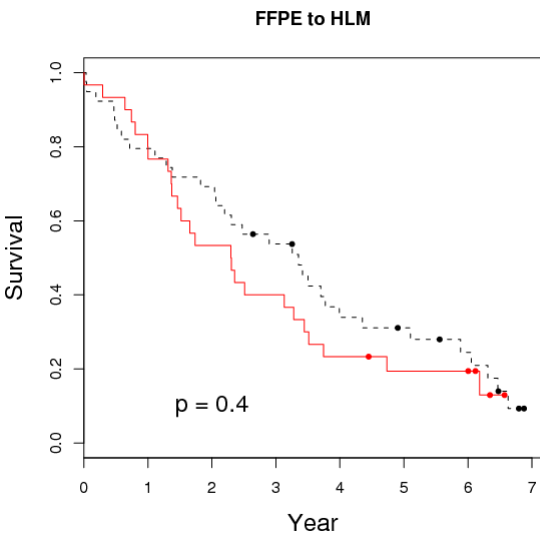
(A)



(B)

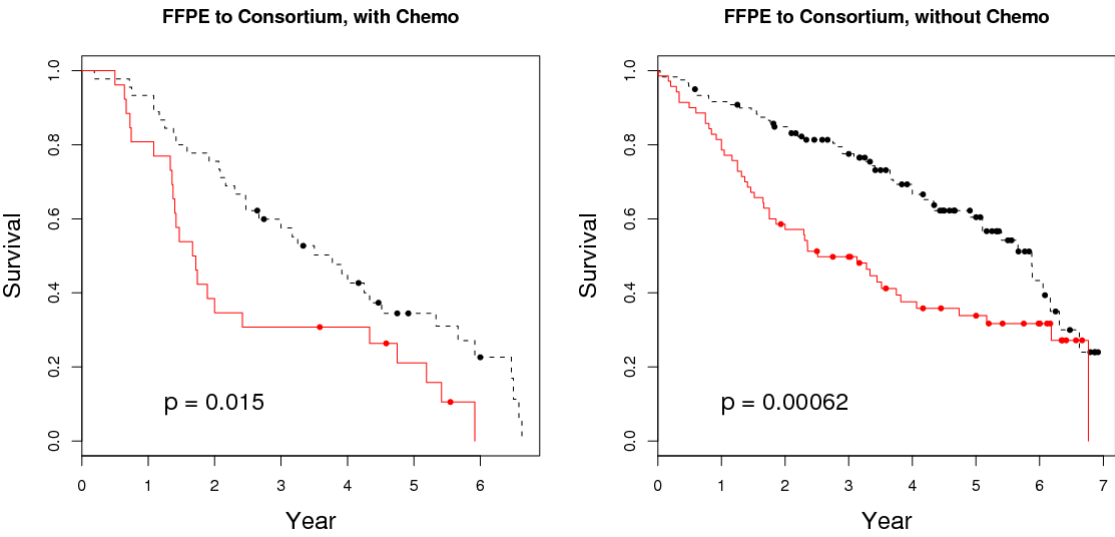


(C)

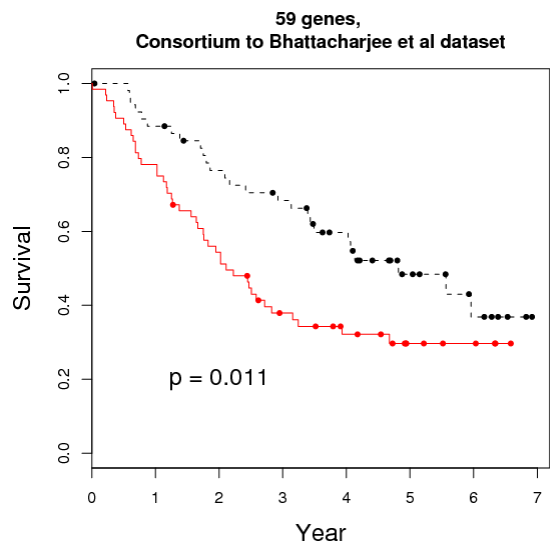


(D)

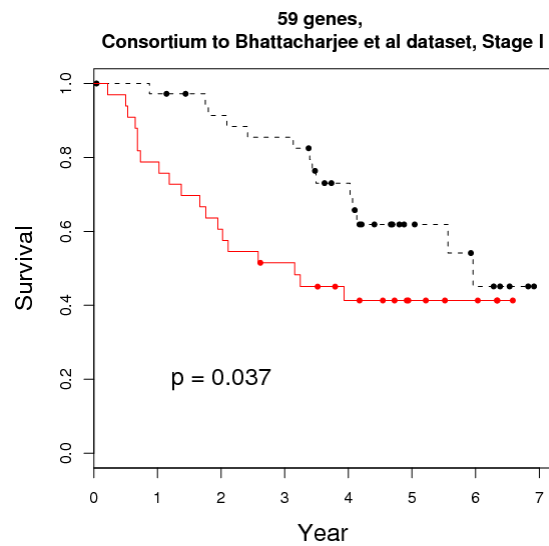
Supplementary Figure 3.



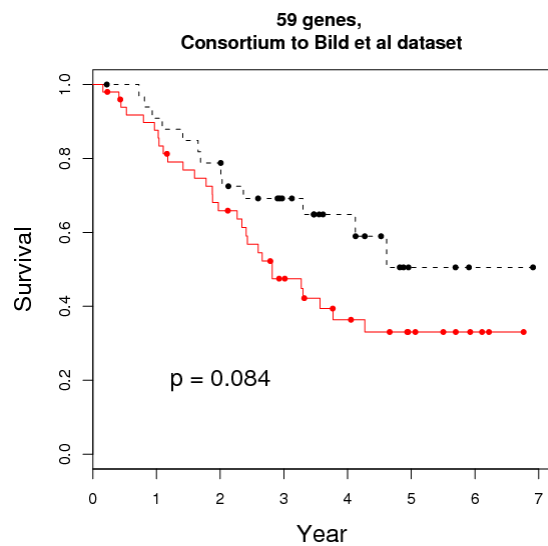
Supplementary Figure 4.



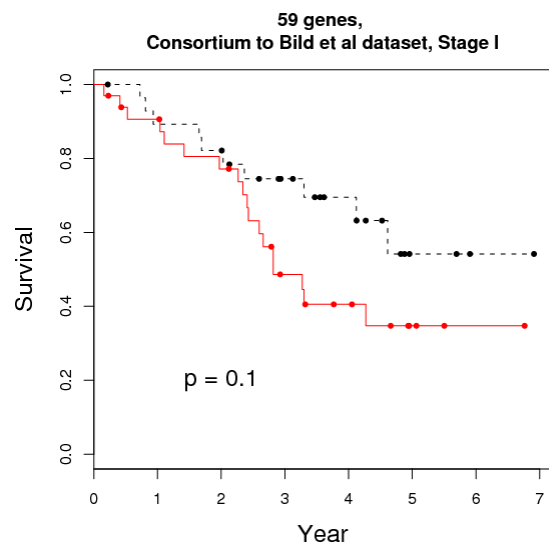
(A)



(B)



(C)



(D)

Table 1 The association between patients' characteristics and RGS risk scores and survival time for consortium patients based on multivariate Cox regression model. (RGS scores were calculated from the prediction model built from MDACC FFPE samples)

Variables	HR (95% CI)	p-value
RGS risk scores	1.300 (1.074, 1.574)	0.0070
Gender (Male vs. Female)	0.803 (0.576, 1.119)	0.1948
Age (continuous in unit of 10 years)	1.571 (1.321, 1.868)	<.0001
Smoking (Current/Former vs. Never)	1.356 (0.791, 2.322)	0.2677
Stage		
Stage II vs. stage I	2.116 (1.433, 3.126)	0.0002
Stage III vs. stage I	4.855 (3.164, 7.449)	<.0001
Adjuvant Chemotherapy (Yes vs. No)	1.688 (1.172, 2.431)	0.0049
Study sites		
DFCI vs. MI	1.295 (0.741, 2.264)	0.3634
HLM vs. MI	1.632 (1.094, 2.434)	0.0163
MSKCC vs. MI	0.657 (0.419, 1.031)	0.0679

Supplementary Table 1 The characteristics of patients from MDACC

Feature	Cohort (n=75)
Age (y)	
Median	65
Range	32 - 82
Mean	64.2
Gender	
Female	27
Male	48
Race	
Caucasian	71
Hispanic	1
Black	2
Asian	1
TNM Stage	
I	34
II	19
III	22
Histology type	
Adenocarcinomas	48
Squamous cell carcinoma	27
Smoking history	
No	7
Yes	68
Adjuvant chemotherapy	
No	28
Yes	47
Follow-up times (y)	
Median	2.76
Range	0.24 – 6.87
Mean	2.79

Abbreviations: TNM, tumor size, node involvement, metastasis status.

Supplementary Table 2. The summary of all the training and testing sets, the P-values of log-rank tests between the predicted high and low risk groups for all the prediction models derived and validated in this study.

Figure	Training	Testing	P-value
3A	FFPE training (N=25)	FFPE testing (N=30)	0.013
3B	Consortium training (N=255)	Consortium testing (N=187)	1.3E-5
3C	FFPE (N = 55)	Consortium (N=442)	5.4E-7
3D	Consortium (N=442)	FFPE (N = 55)	0.068
3E	FFPE (N = 55)	Consortium stage I (N=215)	0.036
3F	FFPE (N = 55)	Consortium stage II (N=82)	0.022
3G	FFPE (N = 55)	Consortium stage III (N=64)	0.021
5A	FFPE (N = 55)	<i>Bhattacharjee et al</i> (N=117)	0.016
5B	FFPE (N = 55)	<i>Bhattacharjee et al</i> stage I (N=70)	0.039
5C	FFPE (N = 55)	<i>Bild et al</i> (N=63)	0.011
5D	FFPE (N = 55)	<i>Bild et al</i> stage I(N=30)	0.024
S2A	FFPE (N = 55)	MSKCC (N= 93)	0.0093
S2B	FFPE (N = 55)	DFCI (N= 64)	0.0076
S2C	FFPE (N = 55)	MI (N= 133)	0.0011
S2D	FFPE (N = 55)	HLM (N= 69)	0.40
S3A	Consortium (N=442)	<i>Bhattacharjee et al</i> (N=117)	0.011
S3B	Consortium (N=442)	<i>Bhattacharjee et al</i> stage I (N=70)	0.037
S3C	Consortium (N=442)	<i>Bild et al</i> (N=63)	0.084
S3D	Consortium (N=442)	<i>Bild et al</i> (N=30)	0.10

Supplementary Table 3 The association between patients' characteristics and RGS groups defined from unsupervised clustering analysis for MDACC patients. P-values for age is based on Wilcoxon test and for other categorical variables are based on Fisher's exact test.

Characteristics	Group 1	Group 2	p-value
Age (median)	65	65	0.47
Male	21 (75%)	17 (63%)	0.57
Current Smoker	17 (61%)	12 (44%)	0.22
Stage			
I	10 (36%)	14 (52%)	0.51
II	8 (28%)	6 (22%)	
III	10 (36%)	7 (26%)	
Histology			
Adenocarcinoma	5 (18%)	25 (93%)	1.3E-8
Squamous	23 (82%)	2 (7%)	

Supplementary Table 4. The top gene sets enriched for RGS group 1 and group2 based on gene set enrichment analysis.

Enriched in group 1

	GS follow link to MSigDB	GS DETAILS	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
1	BRCA_ER_NEG	Details...	82	0.55	2.12	0.000	0.003	0.003	243	tags=49%, list=20%, signal=57%
2	PENG_LEUCINE_DN	Details...	23	0.67	2.00	0.000	0.007	0.015	240	tags=52%, list=19%, signal=63%
3	LI_FETAL_VS_WT_KIDNEY_DN	Details...	15	0.76	1.99	0.000	0.005	0.016	172	tags=67%, list=14%, signal=76%
4	PRMT5_KD_UP	Details...	31	0.62	1.94	0.000	0.009	0.036	159	tags=42%, list=13%, signal=47%
5	HOFFMANN_BIVSBIL_BI_TABLE2	Details...	19	0.66	1.87	0.000	0.017	0.085	161	tags=47%, list=13%, signal=54%
6	ELONGINA_KO_DN	Details...	28	0.61	1.85	0.000	0.019	0.112	141	tags=39%, list=11%, signal=43%
7	TARTE_PLASMA_BLASTIC	Details...	50	0.52	1.82	0.000	0.027	0.177	338	tags=56%, list=27%, signal=74%
8	CIS_XPC_UP	Details...	16	0.66	1.80	0.000	0.030	0.225	81	tags=38%, list=7%, signal=40%
9	PENG_GLUTAMINE_DN	Details...	39	0.54	1.79	0.000	0.028	0.233	257	tags=46%, list=21%, signal=56%
10	HSA04110_CELL_CYCLE	Details...	15	0.68	1.77	0.002	0.032	0.286	252	tags=53%, list=20%, signal=66%
11	STEMCELL_NEURAL_UP	Details...	314	0.38	1.73	0.000	0.044	0.408	344	tags=40%, list=28%, signal=41%
12	STEMCELL_EMBRYONIC_UP	Details...	226	0.39	1.72	0.000	0.046	0.449	313	tags=36%, list=25%, signal=40%
13	HCC_SURVIVAL_GOOD_VS_POOR_DN	Details...	40	0.49	1.62	0.006	0.107	0.778	275	tags=43%, list=22%, signal=53%
14	SERUM_FIBROBLAST_CORE_UP	Details...	22	0.54	1.60	0.021	0.119	0.838	243	tags=55%, list=20%, signal=67%
15	MOREAUX_TACI_HI_VS_LOW_DN	Details...	24	0.54	1.57	0.029	0.146	0.904	252	tags=50%, list=20%, signal=61%

Enriched in group 2

	GS follow link to MSigDB	GS DETAILS	SIZE	ES	NES	NOM p-val	FDR q-val	FWER p-val	RANK AT MAX	LEADING EDGE
1	LI_FETAL_VS_WT_KIDNEY_UP	Details...	24	-0.74	-2.25	0.000	0.000	0.000	138	tags=46%, list=11%, signal=51%
2	BRCA_ER_POS	Details...	53	-0.53	-1.94	0.000	0.035	0.062	197	tags=40%, list=16%, signal=45%
3	TAKEDA_NUP8_HOXA9_8D_DN	Details...	19	-0.65	-1.88	0.004	0.054	0.140	116	tags=37%, list=9%, signal=40%
4	SERUM_FIBROBLAST_CORE_DN	Details...	27	-0.57	-1.73	0.012	0.173	0.482	188	tags=41%, list=15%, signal=47%
5	PENG_RAPAMYCIN_UP	Details...	19	-0.60	-1.71	0.006	0.168	0.548	388	tags=68%, list=31%, signal=98%
6	UEDA_MOUSE_LIVER	Details...	15	-0.63	-1.70	0.013	0.153	0.581	310	tags=53%, list=25%, signal=70%
7	AGEING_KIDNEY_UP	Details...	42	-0.48	-1.64	0.000	0.210	0.742	181	tags=31%, list=15%, signal=35%
8	HSC_MATURE_FETAL	Details...	38	-0.47	-1.60	0.019	0.266	0.860	210	tags=29%, list=17%, signal=34%
9	IRITANI_ADPROX_VASC	Details...	16	-0.59	-1.59	0.036	0.243	0.869	242	tags=50%, list=19%, signal=61%
10	TAKEDA_NUP8_HOXA9_3D_UP	Details...	20	-0.54	-1.59	0.029	0.226	0.874	121	tags=30%, list=10%, signal=33%
11	HSC_MATURE_SHARED	Details...	31	-0.49	-1.59	0.013	0.211	0.882	208	tags=29%, list=17%, signal=34%
12	TPA_SENS_MIDDLE_DN	Details...	26	-0.51	-1.57	0.022	0.213	0.904	235	tags=42%, list=19%, signal=51%

Supplementary Table 5. 59-genes and the association with survival in both FFPE and Consortium set.

Acession	Symbol	Hazard Ratio from FFPE set	p-value from FFPE set	Hazard Ratio from consortium set	p-value from consortium set
NM_002107	H3F3A	0.30	0.034	0.57	0.019
NM_031263	HNRNPK	3.23	0.018	1.73	0.019
NM_002156	HSPD1	2.23	0.010	1.82	0.000
NM_003472	DEK	3.89	0.005	1.43	0.004
NM_016587	CBX3	3.17	0.002	1.78	0.000
NM_173638	NBPF15	0.19	0.008	0.53	0.001
NM_001677	ATP1B1	0.61	0.023	0.87	0.028
NM_005003	NDUFAB1	1.99	0.017	1.57	0.008
NM_005124	NUP153	4.93	0.012	1.52	0.006
NM_004390	CTSH	0.57	0.017	0.69	0.000
NM_014736	KIAA0101	2.04	0.011	1.47	0.000
NM_000935	PLOD2	1.82	0.019	1.22	0.000
NM_012322	LSM5	4.64	0.000	1.90	0.000
NM_002485	NBN	3.39	0.030	1.45	0.012
NM_002453	MTIF2	2.11	0.046	1.42	0.021
NM_002789	PSMA4	3.34	0.022	1.31	0.049
NM_004607	TBCA	3.56	0.026	1.60	0.004
NM_006660	CLPX	4.05	0.017	1.74	0.000
NM_002137	HNRNPA2B1	2.25	0.023	1.83	0.009
NM_001918	DBT	0.25	0.031	0.65	0.001
NM_003096	SNRPG	2.35	0.019	1.67	0.000
NM_003090	SNRPA1	1.95	0.035	1.52	0.003
NM_030881	DDX17	0.17	0.001	0.48	0.000
NM_007208	MRPL3	2.43	0.015	1.42	0.006
NM_002129	HMGB2	1.89	0.025	1.33	0.001
NM_018947	CYCS	2.58	0.027	2.04	0.000
NM_005596	NFIB	0.49	0.013	0.75	0.001
NM_007100	ATP5I	2.20	0.037	1.52	0.016
NM_015149	RGL1	0.44	0.023	0.69	0.000
NM_170662	CBLB	0.42	0.004	0.79	0.036
AL136621	ZMYM2	0.44	0.009	0.56	0.016
NM_006082	TUBA1B	2.44	0.047	1.63	0.005
NM_000712	BLVRA	0.32	0.040	0.73	0.017
NM_033551	LARP1	4.72	0.018	1.82	0.001
NM_015335	MED13L	0.41	0.036	0.78	0.037
AK057191	IDS	0.42	0.032	0.77	0.015
NM_002076	GNS	0.44	0.017	0.68	0.009
NM_005177	ATP6V0A1	0.27	0.015	0.60	0.001
NM_015962	FCF1	3.18	0.007	1.34	0.043
NM_033450	ABCC10	0.28	0.025	0.66	0.001
NM_198843	SFTPB	0.68	0.013	0.85	0.000
NM_001037637	BTF3	2.74	0.047	1.39	0.020
NM_148923	CYB5A	0.49	0.010	0.79	0.000
NM_016061	YPEL5	0.23	0.013	0.66	0.004
NM_016021	UBE2J1	0.26	0.041	0.76	0.044
NM_014056	HIGD1A	3.24	0.009	1.36	0.026
NM_016359	NUSAP1	2.00	0.037	1.31	0.000
NM_021825	CCDC90B	3.17	0.018	1.28	0.010
NM_014167	CCDC59	2.55	0.039	1.67	0.000
NM_013341	OLA1	5.07	0.001	1.58	0.000
NM_030793	FBXO38	0.39	0.021	0.72	0.011
NM_003677	DENR	3.55	0.011	1.34	0.041
NM_144567	ANGEL2	0.18	0.004	0.69	0.011
CR601845	N4BP2L2	0.40	0.009	0.73	0.003
NM_182746	MCM4	2.72	0.001	1.37	0.000
NM_001991	EZH1	0.31	0.034	0.45	0.000
NM_015349	KIAA0240	0.57	0.049	0.54	0.000

NM_001040455	SIDT2	0.41	0.015	0.74	0.031
NM_001012339	DNAJC21	0.20	0.034	0.27	0.001

Supplementary Table 5. The summary of all the training and testing sets, the P-values of log-rank tests between the predicted high and low risk groups for all the prediction models derived and validated in this study.

Figure	Training	Testing	P-value
3A	FFPE training (N=25)	FFPE testing (N=30)	0.013
3B	Consortium training (N=255)	Consortium testing (N=187)	1.3E-5
3C	FFPE (N = 55)	Consortium (N=442)	5.4E-7
3D	Consortium (N=442)	FFPE (N = 55)	0.068
3E	FFPE (N = 55)	Consortium stage I (N=215)	0.036
3F	FFPE (N = 55)	Consortium stage II (N=82)	0.022
3G	FFPE (N = 55)	Consortium stage III (N=64)	0.021
5A	FFPE (N = 55)	<i>Bhattacharjee et al</i> (N=117)	0.016
5B	FFPE (N = 55)	<i>Bhattacharjee et al</i> stage I (N=70)	0.039
5C	FFPE (N = 55)	<i>Bild et al</i> (N=63)	0.011
5D	FFPE (N = 55)	<i>Bild et al</i> stage I(N=30)	0.024
S2A	FFPE (N = 55)	MSKCC (N= 93)	0.0093
S2B	FFPE (N = 55)	DHCI (N= 64)	0.0076
S2C	FFPE (N = 55)	MI (N= 133)	0.0011
S2D	FFPE (N = 55)	HLM (N= 69)	0.40
S3A	Consortium (N=442)	<i>Bhattacharjee et al</i> (N=117)	0.011
S3B	Consortium (N=442)	<i>Bhattacharjee et al</i> stage I (N=70)	0.037
S3C	Consortium (N=442)	<i>Bild et al</i> (N=63)	0.084
S3D	Consortium (N=442)	<i>Bild et al</i> (N=30)	0.10

**VEGFR2 Gene Copy Gain Is Predictive of Shorter Overall Survival in
NSCLC Patients Treated with Adjuvant Chemotherapy**

Fei Yang,^{1,9*} Ximing Tang,^{2*} Carmen Behrens,² Marileila Varella-Garcia,⁵ Lauren A. Byers,² Monique B. Nilsson,² Heather Y. Lin,³ Jing Wang,⁴ Maria Gabriela Raso,¹ J. Jack Lee,³ Luc Girard,⁶ Roy S. Herbst,² John D. Minna,^{6,7,8} John V. Heymach,² and Ignacio I. Wistuba^{1,2}

* These authors contributed equally to this work.

Departments of ¹Pathology, ²Thoracic/Head and Neck Medical Oncology, ³Biostatistics, and ⁴Bioinformatics and Computational Biology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas; ⁵Department of Medicine/Medical Oncology and Pathology, University of Colorado Cancer Center, Aurora, Colorado; ⁶Hamon Center for Therapeutic Oncology and the Departments of ⁷Internal Medicine and ⁸Pharmacology, The University of Texas Southwestern Medical Center, Dallas, Texas; and ⁹Department of Pathology, College of Basic Medicine, Wuhan University, Wuhan, P.R. of China.

† To whom correspondence should be addressed: Ignacio I. Wistuba, MD, Departments of Pathology and Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Tel: 713-563-9184, Fax: 713-563-1848, Email: iiwistuba@mdanderson.org.

Running head: VEGFR2 Gene Copy Gain in NSCLC

Grant Support: This study was supported by grants from the Department of Defense (W81XWH-07-1-0306 W.K.H., J.D.M., J.V.H., I.I.W.), the Specialized Program of Research Excellence in Lung Cancer (P50CA70907 to J.D.M., J.V.H., and I.I.W.; P50CA58187 to MVG), and the National Cancer Institute (Cancer Center Support Grant CA-16672).

Abbreviations: VEGFR2, vascular endothelial growth factor receptor 2; KDR, kinase insert domain receptor; SNP, single nucleotide polymorphism; VEGF, vascular endothelial growth factor; NSCLC, non-small cell lung cancer; RPPA, reverse-phase protein array.

ABSTRACT

Purpose. We studied the role of vascular endothelial growth factor-2 (VEGFR2) gene (*KDR*) abnormalities in malignant cells of surgically resected non-small cell lung carcinoma (NSCLC) tissues and correlated with patients' outcome after treatment with platinum adjuvant chemotherapy.

Patients and Methods. We studied tissues obtained from 248 surgically resected NSCLCs. *KDR* copy number gain (CNG) was examined by quantitative PCR and fluorescence in situ hybridization. VEGFR2 expression and microvascular density were studied by immunohistochemistry. VEGFR2 levels were quantified in 63 NSCLC cell lines by reverse phase protein array and correlated with *in vitro* sensitivity to cisplatin and carboplatin. *KDR* mutation (exons 7, 11 and 21) and single nucleotide polymorphisms (SNPs) 889G/A, 1416A/T and -37A/G were genotyped by PCR-based sequencing.

Results. Malignant cells demonstrated *KDR* CNG in 32% of NSCLC tumors. *KDR* CNG in malignant cells was associated with poor overall survival (OS) (HR=4.0; $P=0.001$) and worse recurrence-free survival (HR=1.83; $P=0.044$) in multivariate analysis. *KDR* CNG predicted worse OS (HR=5.16; $P=0.003$) in patients who received platinum adjuvant therapy but not in untreated patients ($P=0.349$). Higher VEGFR2 expression levels in cell lines significantly correlated with resistance to cisplatin ($P=0.0005$) and carboplatin ($P=0.005$). *KDR* mutations were not detected in NSCLC tumor tissues. The *KDR* variant genotypes SNPs 1416 AT/TT and -37 AG/GG were associated with a favorable OS in lung adenocarcinoma.

Conclusion. The association between *KDR* CNG and worse outcome in platinum adjuvant therapy-treated NSCLC patients suggests that might be a potential biomarker for predicting the efficacy of adjuvant chemotherapy in this disease.

INTRODUCTION

Tumor growth is critically dependent on neovascularization, which is a well-established aspect of cancer biology.¹ The ligand vascular endothelial growth factor (VEGF) is an endothelial cell-specific mitogen known to be a highly potent and specific mediator of angiogenesis, and has two identified receptors.²⁻⁴ The VEGF receptor 2 (VEGFR2) coded by the gene *KDR* (located in 4q12) is the predominant mediator of VEGF-stimulated endothelial cell functions, including cell migration, proliferation, survival, and enhancement of vascular permeability.^{5,6} VEGFR2 exhibits robust protein-tyrosine kinase activity in response to the VEGF ligand.⁷

In human epithelial tumors, including lung, VEGFR2 has shown to be expressed in malignant cells as well as in the endothelial cell of tumor vasculature.⁸⁻¹¹ In non-small cell lung carcinoma (NSCLC), VEGFR2 has been found to be overexpressed in malignant cells of tumor tissues, and associated with a poor outcome.⁸⁻¹¹ Recently, a relatively high frequency (9%) of mutation and amplification of *KDR* has been detected in lung adenocarcinoma histology.¹² however, the presence of these abnormalities in squamous cell carcinomas of the lung is unknown. In addition, there is no data available on the correlation of *KDR* abnormalities with tumor and patients' characteristics in lung cancer, including outcome and response to therapy.

The objective of this study was to characterize the molecular abnormalities of VEGFR2 in epithelial malignant cells of NSCLC major histology types adenocarcinoma and squamous cell carcinoma, and correlate with patients'

clinical characteristics. We studied *KDR* copy number gain (CNG), mutation and genetic variations in malignant cells of surgically resected NSCLC tumor tissues and correlated with tumor's pathological features and with patient's platinum adjuvant treatment and outcome.

MATERIALS AND METHODS

Case Selection. We obtained archived frozen and formalin-fixed and paraffin-embedded (FFPE) tissues from NSCLC surgically resected with curative intent from the Lung Cancer Specialized Program of Research Excellence (SPORE) tissue bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX). The tissue banking and the study were approved by the Institutional Review Board. The tumors were classified using the 2004 World Health Organization (WHO) classification system.¹³ We randomly selected 248 NSCLC specimens (159 adenocarcinomas and 89 squamous cell carcinomas) to test *KDR* abnormalities. Detailed clinical and pathologic information of the cases is presented in Table 1. The median follow-up of the patients was 3.53 years for those who are censored.

DNA Extraction. To enrich for malignant cell content, tumor tissues were manually microdissected from optimal cutting temperature (OCT) compound-embedded frozen tissue sections for subsequent DNA extraction. Tumor DNA was extracted using Pico Pure DNA Extraction Kit (Arcturus, Mountain View, CA) according to the manufacturer's instructions.

Real-time Quantitative PCR (qPCR) for *KDR* Copy Number Analysis. DNA samples from tumor with proportions of microdissected tumor cell greater than 70% were qualified for qPCR analysis. *KDR* gene copy number was detected by real-time quantitative PCR (qPCR) using the ABI 7300 real time PCR system

(Applied Biosystems, Foster City, CA). The primers used to amplify *KDR* were as follows: KF-GACACACCCTCAGGCTCTTG, and KR-
 ACTTTTCACCGCCTGTTCTC. Each PCR was performed using *Power* SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) at 50°C for 2 min and 95°C for 10 min followed by 40 cycles at 95°C for 15 s and 60°C for 1 min. β -*Actin* was introduced as the endogenous reference gene and TaqMan Control Human Genomic DNA (Applied Biosystems, Foster City, CA) was amplified as a standard control for calibration. All sample and standard DNA reactions were set in triplicate to gauge reaction accuracy. The target gene copy number was quantified using the comparative C_t method. Gene copy number of greater than 4 was considered as increased, as previously reported.¹⁴

***KDR* Fluorescence *in Situ* Hybridization (FISH) Analysis.** We analyzed the gene copy number per cell using a dual-color FISH assay developed by one of the co-authors (M. V-G). The *KDR* probe was prepared from the BAC clone RP11-21A18 obtained from CHORI (Oakland, CA). Detailed information on the FISH assay is provided in the Supplementary Methods. Copy number analysis was done in approximately 50 malignant cells nuclei per tumor in at least four areas. Greater than 2 copies of the gene per cell on average was considered as copy number gain.

NSCLC Cell Lines VEGFR2 Analysis by Reverse-Phase Protein Array (RPPA). Sixty-three NSCLC cell lines were provided by a co-author of this study

(J.D.M). The identity of each cell line was confirmed by DNA fingerprinting. Cisplatin and carboplatin sensitivity was determined by MTS assay for each cell line and the concentration required for 50% growth inhibition (IC₅₀) was determined. MTS assays were repeated at least three times for each cell line and the mean IC₅₀ value used for analysis. For RPPA analysis, protein lysate was collected from subconfluent cultures after 24 hours growth in media with 10% fetal bovine serum (FBS) and assayed by RPPA as previously described.^{15,16}

VEGFR2 Expression and Microvascular Density (MVD) Analyses. Histology tissue sections were incubated with primary antibodies against VEGFR2 (dilution 1:50, rabbit polyclonal antibody, Abcam, Cambridge, MA) for 90 min, and CD34 (dilution 1:100, mouse monoclonal antibody, Lab Vision, Fremont, CA) for 35 min at room temperature. Tissue sections were then incubated with the secondary antibody (EnVision Dual Link+; DAKO, Carpinteria, CA) for 30 min, after which diaminobenzidine chromogen was applied for 5 min.

Protein expression was quantified immunohistochemically using light microscopy with an original ×20 magnification by two observers (F. Y. and I. W.). Immunostained samples of the cytoplasm and membranes were analyzed for VEGFR2 using a 4-value intensity score (0, 1+, 2+, 3+) and the percentage (0% to 100%) of the extent of reactivity. The final score was obtained by multiplying the intensity and extent-of-reactivity values (range, 0 to 300). MVD was assessed

by Ariol^R 2.0 Image System (Ariol^R, Applied Imaging Inc, San Jose, CA) using the criteria of Weidner *et al.*¹⁷

KDR Mutation and SNPs Genotyping Analyses. For *KDR* mutation analysis in NSCLC cell lines we examined exons 7, 11, 21, 26, 27 and 30, using PCR-based sequencing and intron-based PCR primers (details in Supplementary Methods). NSCLC tumor specimens were examined for *KDR* exons 7, 11 and 21 for the identification of gene mutations, and for the genotyping of the three SNPs that we identified in these exons: 889G/A, rs2305948 (exon 7); 1416 A/T, rs1870377 (exon 11); and, -37A/G, rs2219471 (exon 21). Mutation and SNP genotyping were performed using the ABI Prism 7900 Sequence Detection System (Applied Biosystems, Foster City, CA).

Statistical Analysis. Data were summarized using descriptive statistics such as frequency distribution, mean (\pm s.d.) and median (range) accompanied by graphs. Demographic and clinical information was compared using the Chi-squared or Fisher exact tests (for category variables) and Wilcoxon rank-sum or Kruskal-Wallis tests (for continuous variables). The distributions of overall survival (OS) and recurrence-free survival (RFS) were estimated by the Kaplan-Meier method and compared between groups using the log-rank test. Cox proportional hazard models were used for regression analyses of survival data and conducted on OS defined as time from surgery to death or last contact, and on RFS defined as time from surgery to recurrence or last contact. Follow-up time was censored at 5

years. The NSCLC cell lines RPPA data was quantified using the SuperCurve method which detects changes in protein level as previously reported.¹⁸

RESULTS

***KDR* Gene Copy Gain Analysis.** We examined DNA extracted from 37 NSCLC cell lines part of the UT-Lung Cancer SPORE cell line repository, provided by one of the co-authors (J. D. M.). Using qPCR, two cancer cell lines H661 and H889 showed *KDR* gene copy gain (≥ 4 copies); these results were confirmed by the available aCGH results (data not shown).

In epithelial malignant NSCLC cells microdissected from tumor tissues, *KDR* copy gains were detected in 45 (32%) of 139 tumors examined. Similar frequency of *KDR* copy gain was found in adenocarcinoma (26/85, 31%) and squamous cell carcinoma (19/54, 35%) histologies ($P=0.572$). The range of increased *KDR* copy numbers was from 4.0 to 11.0 gene copies. None of 15 normal tissue samples adjacent to the NSCLC tested showed *KDR* gene copy gain.

To confirm *KDR* copy gain results by qPCR, 30 tumor specimens were selected for FISH analysis, including 20 cases with and 10 cases without *KDR* copy gain by qPCR. *KDR* copy gains was confirmed in the malignant cells in all 20 NSCLC specimens detected by qPCR, and also detected in 6 out of 10 cases without copy number gain by qPCR. This correlation was statistically significant (Kappa coefficient [95% CI]=0.471, 0.149 to 0.792; $P=0.007$) (Figure 1, Panel A). The copy numbers of *KDR* by FISH and qPCR were significantly correlated ($r=0.417$, $P=0.022$), and concordance was observed in 24 out of 30 (80%) cases.

Correlation between *KDR* Gene Copy Gain and VEGFR2 Protein Expression

and MVD. To assess the immunohistochemical expression of VEGFR2 in NSCLC malignant cells and the MVD (CD34) in lung tumor tissue stroma, we selected 52 lung tumor specimens with whole histologic sections from FFPE tissues (Figure 1, Panels B and C). Of these, 26 cases had *KDR* copy number gain and 26 cases did not. VEGFR2 protein expression was present both in the cytoplasm and membranes of malignant cells as well as in vessel endothelial cells (Figure 1, Panel B).

Levels of VEGFR2 expression in cytoplasm and in membrane were associated with *KDR* gene copy gain in malignant cells of NSCLC. Tumors with *KDR* gene copy gains showed significantly higher cytoplasmic ($P=0.013$) and membrane ($P=0.009$) VEGFR2 protein expression in the malignant cells, and higher MVD ($P=0.018$) and larger vessel areas ($P=0.033$) in the tumor stroma than cases without *KDR* gene copy gain (Figure 2).

Association between *KDR* Copy Number Gain and NSCLC

Clinicopathologic Features. When we correlated *KDR* copy number gain with patients' clinicopathologic features, we did not find correlation with tumor histology, smoking status, and tumor stage. The platinum neoadjuvant-treated tumors (33%, 8/24) had similar frequency of *KDR* copy number gains than cases without neoadjuvant therapy (32%, 37/115; $P=0.912$).

Interestingly, *KDR* copy gain was associated with poor OS (HR=4.0; 95% CI, 1.76 to 9.07; $P=0.001$) and shorten RFS (HR=1.83, 95% CI, 1.02 to 3.29;

$P=0.044$) in 115 NSCLC patients who have not received neoadjuvant therapy in the multivariate analysis after adjusting for stage and adjuvant therapy. Strikingly, *KDR* copy gain predicted a worse OS (HR=5.16, 95% CI, 1.75 to 15.2, $P=0.003$) in NSCLC patients receiving platinum adjuvant therapy but did not predict a poor OS ($P=0.349$) in patients without adjuvant therapy (Figure 3 and Table 2). These data suggest that *KDR* CNG in malignant cells may represent a predictive marker of worse outcome in patients with surgically resected NSCLC treated with adjuvant chemotherapy.

VEGFR2 Protein Levels and Correlation with Platinum Resistance in

NSCLC Cell Lines. The association detected between *KDR* CNG and worse outcome in patients treated with platinum adjuvant therapy prompted us to examine the correlation between *KDR* CNG and VEGFR2 protein levels and NSCLC cell lines *in vitro* resistance to platinum-based chemotherapy drugs. We did not find association between *KDR* CNG by aCGH in NSCLC cell lines (n=37) with *in vitro* resistance to cisplatin and carboplatin (data not shown). However, when we correlated the expression of VEGFR2 protein in a panel of 63 untreated NSCLC cell lines by RPPA with each cell line's sensitivity to cisplatin or carboplatin we found that higher VEGFR2 expression levels were significantly associated with resistance to both cisplatin and carboplatin by Pearson correlation (Supplementary Figure 1). The correlation coefficient between VEGFR2 expression and the concentration of cisplatin required to inhibit cell

growth by 50% (IC50) was $r=0.346$ ($P=0.005$). Similarly, carboplatin IC50s were correlated to VEGFR2 levels with an $r=0.319$ ($P=0.011$).

KDR Mutation Analysis. For *KDR* mutation analysis in NSCLC cell lines, we examined six *KDR* exons (7, 11, 21, 26, 27 and 30) shown to be mutant in adenocarcinoma tumors in a study published by Ding *et al.*¹² In 37 tested NSCLC cell lines we found only two mutations in the *KDR* gene, an intronic T+2A exon 11 mutation in HCC2450, and a CGT946CAT point mutation in exon 21 in HCC2279. No mutation affecting exons 11 and 21 were detected in 200 NSCLC tissues specimens examined.

KDR SNP Analysis. Three *KDR* SNPs (889G/A, 1416A/T, and -37A/G) were genotyped in DNA extracted from 200 NSCLC tumors (Supplementary Table 1), and correlated with patients clinicopathologic features, including outcome. We did not find correlation between the SNP genotypes distribution and OS and RFS of all NSCLC patients examined. When we analyzed the data by tumor histology, we found that among the 127 lung adenocarcinoma patients examined both *KDR* 1416 AT/TT (HR=0.45; 95% CI, 0.2 to 0.99; $P=0.048$) and -37 AG/GG (HR=0.43; 95% CI, 0.2 to 0.92; $P=0.031$) variant genotypes were associated with a favorable OS in the multivariate analysis after adjusting for tumor stage and neoadjuvant therapy (Supplementary Figure 2 and Table 3). However, no *KDR* SNP genotype was associated with OS in lung squamous cell carcinoma patients

(Supplementary Figure 2). Moreover, no genotype in the three *KDR* SNPs was associated with RFS in NSCLC patients divided by histology type.

Furthermore, among NSCLC patients with the *KDR* 889 GA/AA variant genotypes, those who received platinum neoadjuvant and/or adjuvant chemotherapy showed a significantly better OS (HR=0.22; 95% CI, 0.05 to 0.94; $P=0.041$) than patients who did not receive chemotherapy in the multivariate analysis after adjusting for histology and tumor stage. However, no survival benefit was found in NSCLC patients with *KDR* 889 GG wild genotype (HR=1.23; 95% CI, 0.64 to 2.35; $P=0.538$).

Finally, all *KDR* SNP genotypes were compared with primary tumor expression for VEGFR2 and MVD in 52 NSCLC specimens. However, no genotypes correlated with the expression of any of these markers in NSCLC tumors.

DISCUSSION

Our study represents the first report in lung cancer showing a relatively high frequency of *KDR* copy number gain (32%) in both major histology types of NSCLC, adenocarcinoma and squamous cell carcinoma, by qPCR and confirmed in a subset of cases by FISH. Conversely, mutations of *KDR* were rarely detected in NSCLC cell lines and not detected in tumor specimens. Notably, *KDR* copy gain was associated with poor survival of patients receiving platinum adjuvant therapy but not in patients without adjuvant therapy. These data suggest that *KDR* CNG in malignant cells may represent a predictive marker of worse outcome in patients with surgically resected NSCLC treated with platinum adjuvant chemotherapy.

In our study, tumors with *KDR* gene copy gain in the malignant cells showed significantly higher VEGFR2 protein expression in the cytoplasm and membrane of those cells, as well as higher MVD and larger vessel areas in the tumor stroma, compared with tumors lacking the *KDR* gene copy gain. Therefore, *KDR* gene copy gain might result in the overexpression of VEGFR2 in lung cancer cells, and subsequently, may promote an enhancement of tumor blood vessels formation. One possible explanation for this association is that VEGFR2-overexpressing lung cancer cells may intensify the surrounded levels of VEGF which subsequently stimulates tumor tissue angiogenesis resulting in higher MVD.

A provocative finding of this study is the effect of *KDR* copy gain in malignant cells in the outcome of NSCLC patients receiving platinum adjuvant

chemotherapy after surgical resection with curative intent. We report for the first time that *KDR* gene copy gain in tumor cells predicted a worse OS in NSCLC patients receiving platinum-based adjuvant therapy but did not predict a poor OS in patients without adjuvant therapy. These findings suggest that *KDR* copy gain represents a potential biomarker for predicting the efficacy of adjuvant platinum chemotherapy in NSCLC patients. Although the molecular mechanisms of this association are unclear, our findings that higher levels of VEGFR2 expression by RPPA in 63 NSCLC cell lines were significantly correlated with *in vitro* resistance to both cisplatin and carboplatin provides support to our clinical observation. However, further prospective studies with larger patient cohorts are needed to assess the role of *KDR* copy gain in malignant cells and outcome of NSCLC patients treated with platinum-based chemotherapy in both surgically resected and advanced metastatic tumor settings.

In our study, the variant genotypes of *KDR* SNPs 1416 (AT/TT) and -37 (AG/GG) associated with a favorable OS in the multivariate analysis. Ours is the first report showing association between *KDR* SNP genotypes and prognosis in lung cancer. In breast cancer patients the *KDR* SNP 1416 A/T genotypic variant was associated with the expression of progesterone receptors, and its presence suggested a better prognosis for carriers of the T allele.¹⁹ Questions remain about the functional roles of the *KDR* SNPs responsible for the associations with outcome of NSCLC patients, particularly in adenocarcinoma patients, found in our study. Interestingly, the *KDR* SNP 1416 A/T (Q472H), a non-synonymous coding polymorphism, is located in the fifth immunoglobulin-like domain within

the extracellular region of VEGFR2 and is important for preventing VEGF-independent receptor dimerization and signal transduction.²⁰ The other prognostic *KDR* SNP in lung adenocarcinoma patients, SNP -37AG/GG is located in intron 11 within the protein kinase domain and it has not been associated to any specific protein functional effect. However, functional studies are needed to characterize further the importance of the *KDR* SNPs in NSCLC.

In summary, our findings indicate that *KDR* copy number gain was frequently detected in NSCLC tumors and associated with patients' poor survival and tumors' angiogenesis which might be a potential biomarker for predicting the efficacy of platinum adjuvant chemotherapy in NSCLC patients. In addition, *KDR* SNP genotypes correlate with outcome in patients with surgically resected NSCLC tumors. To our knowledge, this is the first report to demonstrate the clinical importance of copy number gain and genetic variations of *KDR* in NSCLC. However, further studies with larger patient cohorts are needed to ascertain the clinical effects of the molecular changes of *KDR* in lung cancer and to elucidate the roles of angiogenesis-related pathways in lung cancer progression and response to adjuvant chemotherapy.

REFERENCES

1. Folkman J: Tumor angiogenesis: therapeutic implications. *N Engl J Med* 285:1182-6, 1971
2. Fidler IJ, Ellis LM: The implications of angiogenesis for the biology and therapy of cancer metastasis. *Cell* 79:185-8, 1994
3. Ferrara N, Davis-Smyth T: The biology of vascular endothelial growth factor. *Endocr Rev* 18:4-25, 1997
4. Hanahan D, Weinberg RA: The hallmarks of cancer. *Cell* 100:57-70, 2000
5. Terman BI, Carrion ME, Kovacs E, et al: Identification of a new endothelial cell growth factor receptor tyrosine kinase. *Oncogene* 6:1677-83, 1991
6. Bernatchez PN, Soker S, Sirois MG: Vascular endothelial growth factor effect on endothelial cell proliferation, migration, and platelet-activating factor synthesis is Flk-1-dependent. *J Biol Chem* 274:31047-54, 1999
7. Waltenberger J, Claesson-Welsh L, Siegbahn A, et al: Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *J Biol Chem* 269:26988-95, 1994
8. Ishii H, Yazawa T, Sato H, et al: Enhancement of pleural dissemination and lymph node metastasis of intrathoracic lung cancer cells by vascular endothelial growth factors (VEGFs). *Lung Cancer* 45:325-37, 2004

9. Ludovini V, Gregorc V, Pistola L, et al: Vascular endothelial growth factor, p53, Rb, Bcl-2 expression and response to chemotherapy in advanced non-small cell lung cancer. *Lung Cancer* 46:77-85, 2004
10. Seto T, Higashiyama M, Funai H, et al: Prognostic value of expression of vascular endothelial growth factor and its flt-1 and KDR receptors in stage I non-small-cell lung cancer. *Lung Cancer* 53:91-6, 2006
11. Carrillo de Santa Pau E, Arias FC, Caso Pelaez E, et al: Prognostic significance of the expression of vascular endothelial growth factors A, B, C, and D and their receptors R1, R2, and R3 in patients with nonsmall cell lung cancer. *Cancer* 115:1701-12, 2009
12. Ding L, Getz G, Wheeler DA, et al: Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069-75, 2008
13. Mountain CF: Revisions in the International System for Staging Lung Cancer. *Chest* 111:1710-7, 1997
14. Yamamoto H, Shigematsu H, Nomura M, et al: PIK3CA mutations and copy number gains in human lung cancers. *Cancer Res* 68:6913-21, 2008
15. Cheng KW, Lu Y, Mills GB: Assay of Rab25 function in ovarian and breast cancers. *Methods Enzymol* 403:202-15, 2005
16. Byers LA, Sen B, Saigal B, et al: Reciprocal regulation of c-Src and STAT3 in non-small cell lung cancer. *Clin Cancer Res* 15:6852-61, 2009
17. Weidner N, Semple JP, Welch WR, et al: Tumor angiogenesis and metastasis--correlation in invasive breast carcinoma. *N Engl J Med* 324:1-8, 1991

18. Hu J, He X, Baggerly KA, et al: Non-parametric quantification of protein lysate arrays. *Bioinformatics* 23:1986-94, 2007
19. Forsti A, Jin Q, Altieri A, et al: Polymorphisms in the KDR and POSTN genes: association with breast cancer susceptibility and prognosis. *Breast Cancer Res Treat* 101:83-93, 2007
20. Tao Q, Backer MV, Backer JM, et al: Kinase insert domain receptor (KDR) extracellular immunoglobulin-like domains 4-7 contain structural features that block receptor dimerization and vascular endothelial growth factor-induced signaling. *J Biol Chem* 276:21916-23, 2001

FIGURE LEGENDS

Figure 1. *Panel A*, representative examples of *KDR* copy number examined by fluorescence in situ hybridization (FISH) in non-small cell carcinoma (NSCLC) tissue specimens. *a*, copy number gain; *b*, no copy number gain. Red signals represent the *KDR* gene probe, and green signals the internal control probe (magnification $\times 1000$). *Panel B*, representative example of immunohistochemical expression of VEGFR2 in NSCLC tissue specimens. VEGFR2 protein expression was present both in the cytoplasm and membranes of tumor cells in adenocarcinoma (*c*) and squamous cell carcinoma (*d*) (magnification $\times 200$). *Panel C*, representative example of immunohistochemical expression of CD34-positive vessels (microvascular density, MVD) in adenocarcinoma (*e*) and squamous cell carcinoma (*f*) (magnification $\times 200$).

Figure 2. Expression of VEGFR2 and microvascular density (MVD) in primary lung cancers with *KDR* copy number gain compared with lung cancers without *KDR* copy number gain. The box-plots depicting scores of immunohistochemical expression of VEGFR2 cytoplasm, VEGFR2 membrane, and MVD and vessel area (mm^2) comparing 26 lung cancers having *KDR* copy number gain ($\text{CNG} \geq 4$) with 26 lung cancers without *KDR* copy number gain ($\text{CNG} < 4$).

Figure 3. Kaplan-Meier curve for overall survival (OS) and recurrence-free survival (RFS) by *KDR* copy number gain in non small cell lung cancer patients

and two subgroups of platinum adjuvant therapy and without adjuvant therapy. All patients included in these analyses did not receive neoadjuvant therapy.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. VEGFR2 protein expression by RPPA associated with *in vitro* resistance to platinum drugs. Correlation between the concentration of (A) cisplatin and (B) carboplatin required to inhibit NSCLC cell growth (IC50) and VEGFR2 levels.

Supplementary Figure 2. Kaplan-Meier curve for overall survival (OS) by genotypes of two *KDR* single nucleotide polymorphisms in adenocarcinoma and squamous cell carcinoma of lung.

Table 1. Clinicopathologic characteristic of non-small cell lung carcinoma examined for *KDR* abnormalities.

Characteristic	All Cases Tested (N=248) Number (%)	Cases Tested for Copy Gain (N=139) Number (%)	Cases Tested for SNPs* (N=200) Number (%)
Mean Age in Years (range)	64.6 (26.4- 86.9)	64.9 (32.2- 84)	63.97 (26.4- 86.9)
Gender			
Female	110 (44)	57 (41)	88 (44)
Male	138 (56)	82 (59)	112 (56)
Tumor Histology			
Adenocarcinoma	159 (64)	85 (61)	127 (64)
Squamous cell carcinoma	89 (36)	54 (39)	73 (36)
TNM Pathology Stage			
I	120 (49)	70 (51)	86 (43)
II	50 (20)	28 (20)	40 (20)
III	72 (29)	39 (28)	68 (34)
IV	6 (2)	2 (1)	6 (3)
Smoking status [‡]			
Current	102(41)	52 (37)	89 (45)
Former	108 (44)	64 (46)	82 (41)
Never	38 (15)	23 (17)	29 (14)
Neoadjuvant therapy [†]			
No	181(73)	115 (83)	133 (67)
Yes	67 (27)	24 (17)	67 (33)
Adjuvant therapy [†]			
No	138 (56)	69 (50)	90 (45)
Yes	110 (44)	70 (50)	110 (55)

* SNP, Single Nucleotide Polymorphism.

[‡] Patients who had smoked at least 100 cigarettes in their lifetime were defined as ever smokers, and smokers who quit smoking at least 12 months before lung cancer diagnosis were defined as former smokers.[†] All patients who received neoadjuvant and adjuvant chemotherapy received platinum (cisplatin or carboplatin), and the chemotherapy regimen most frequently administered was carboplatin-taxol combination.

Table 2. Multivariate analysis for outcome by *KDR* copy gain in non-small cell lung carcinoma patients by adjuvant chemotherapy.

Cases	N	Comparison	Outcome	Adjusted Hazard Ratio (HR) [*] (95% CI)	P
All patients	115	Gain vs. no gain	OS [‡]	4.00 (1.76, 9.07)	0.001
			RFS [†]	1.83 (1.02, 3.29)	0.044
Adjuvant therapy	61	Gain vs. no gain	OS	5.16 (1.75, 15.2)	0.003
			RFS	1.87 (0.9, 3.92)	0.1
No adjuvant therapy	54	Gain vs. no gain	OS	1.99 (0.47, 8.4)	0.349
			RFS	1.83 (0.66, 5.05)	0.243

^{*} Adjusting for tumor stage; follow-up is censored at 5 years.

[‡] OS, overall survival

[†] RFS, recurrence-free survival

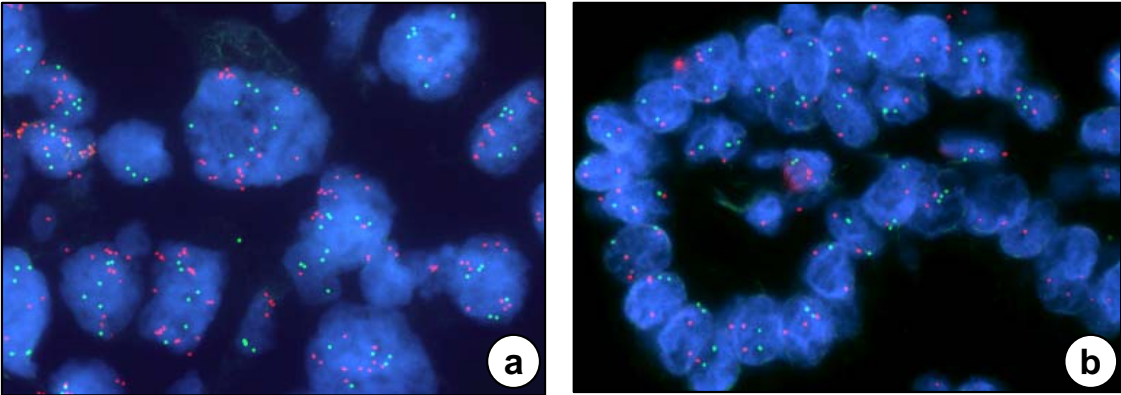
Table 3. Multivariate analysis for overall survival in three *KDR* single nucleotide polymorphisms (SNP) in non-small cell lung carcinoma.

Cases	<i>KDR</i> SNP	Genotype	Adjusted Hazard Ratio (HR)* (95% CI)	<i>P</i>
NSCLC	889	GA/AA vs. GG	0.92 (0.51 to 1.66)	0.78
	1416	AT/TT vs. AA	0.59 (0.34 to 1.01)	0.056
	-37	AG/GG vs. AA	0.6 (0.35 to 1.03)	0.062
Adenocarcinoma	889	GA/AA vs. GG	0.63 (0.24 to 1.65)	0.348
	1416	AT/TT vs. AA	0.45 (0.2 to 0.99)	0.048
	-37	AG/GG vs. AA	0.43 (0.2 to 0.92)	0.031
Squamous cell carcinoma	889	GA/AA vs. GG	1.16 (0.53 to 2.51)	0.713
	1416	AT/TT vs. AA	0.76 (0.36 to 1.61)	0.468
	-37	AG/GG vs. AA	0.84 (0.4 to 1.78)	0.649

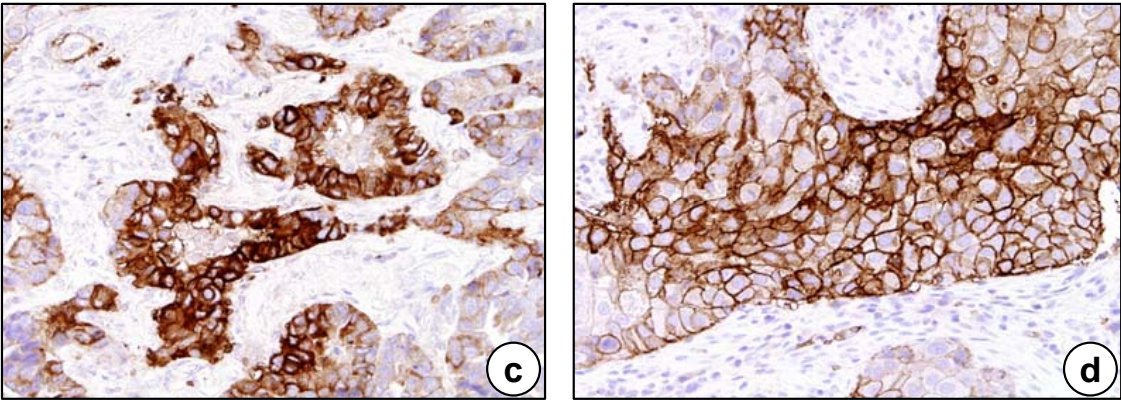
* Adjusting for tumor stage; follow-up is censored at 5 years.

Figure 1

A. FISH - *KDR*



B. Immunohistochemistry - VEGFR2



C. Immunohistochemistry - MVD (CD34)

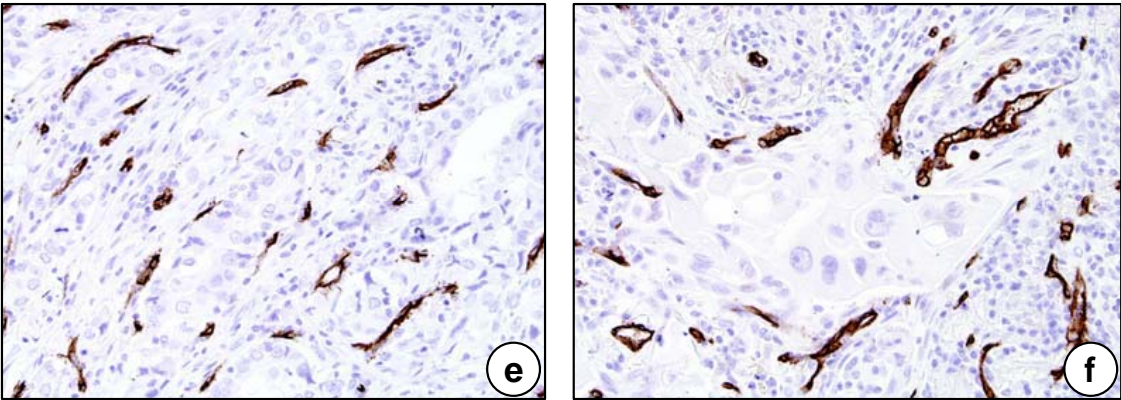


Figure 2

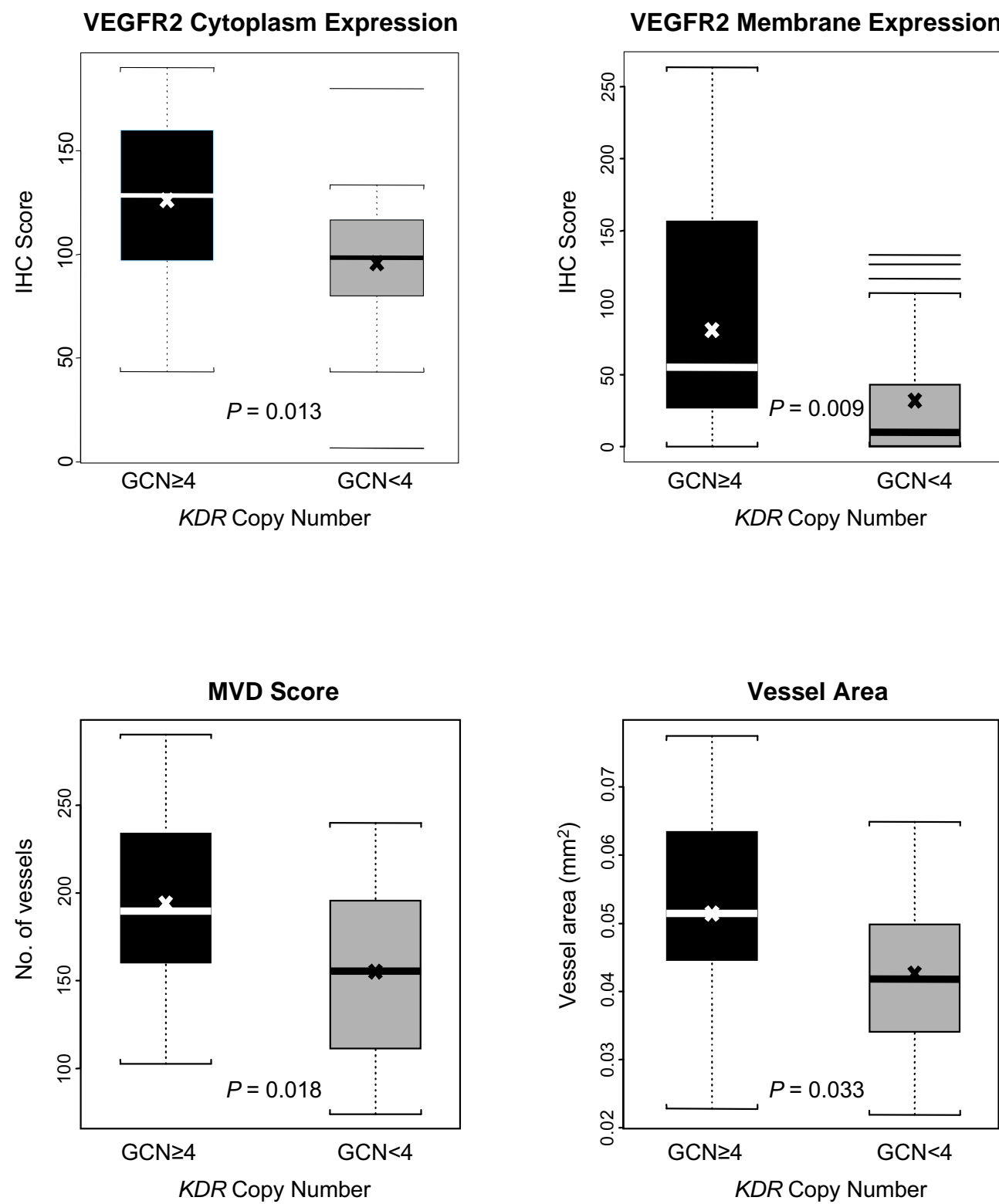
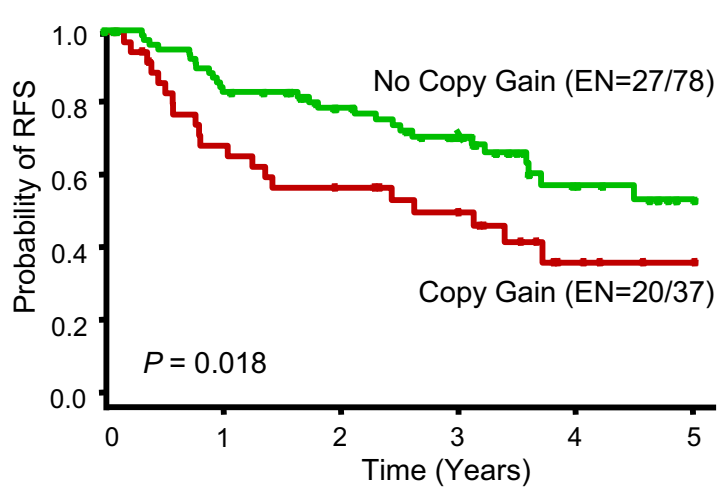
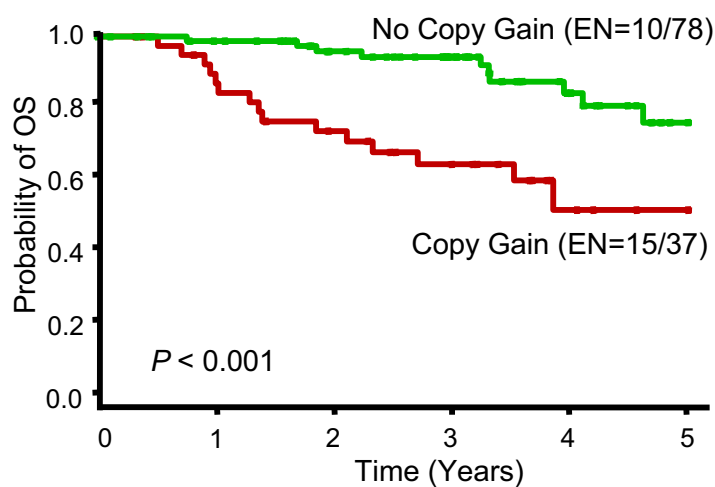


Figure 3

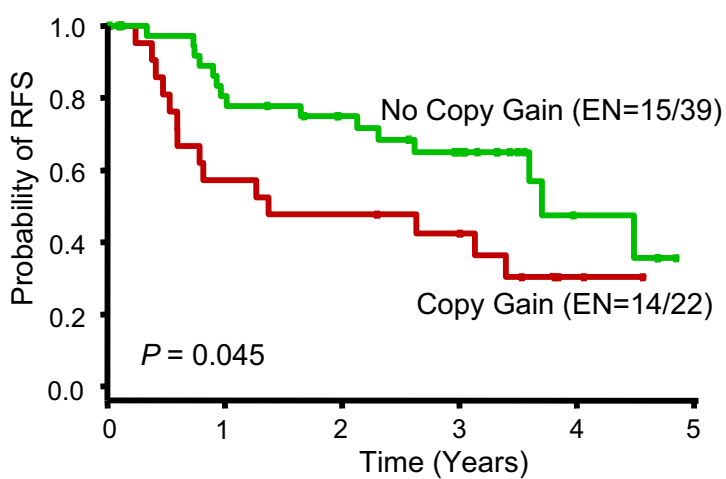
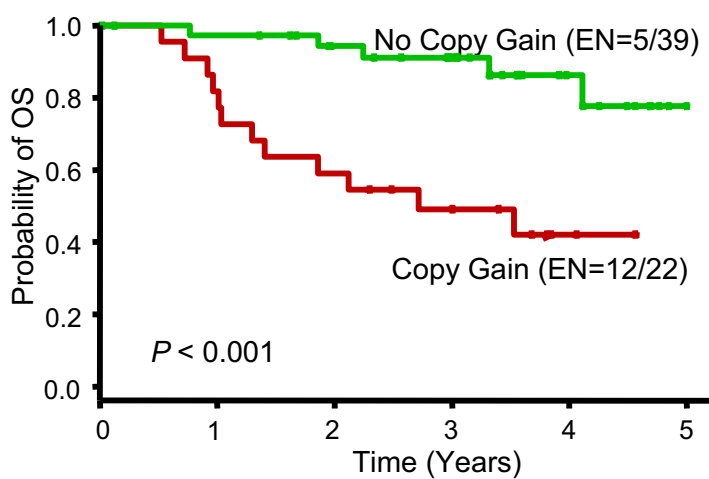
Overall Survival

Recurrence-free Survival

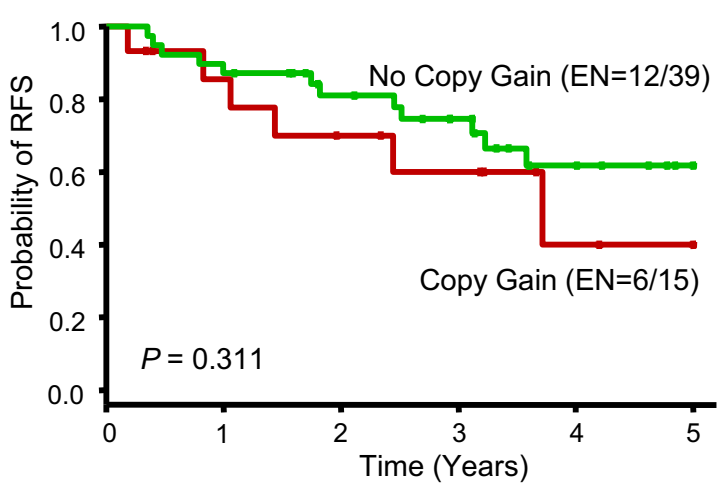
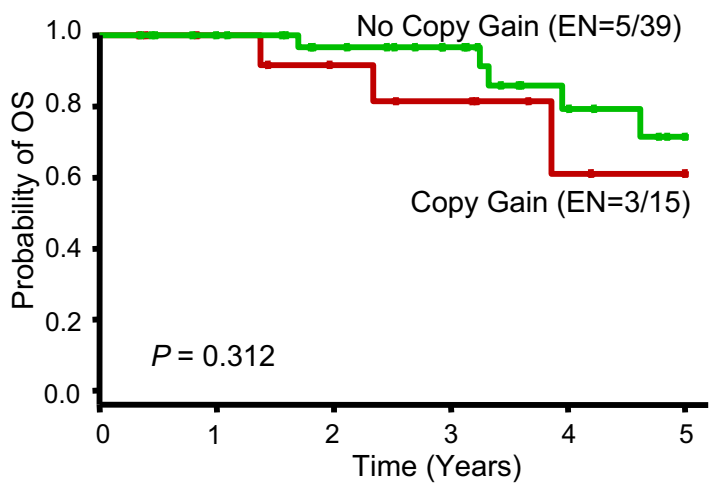
NSCLC (All Patients)



NSCLC with Adjuvant Therapy



NSCLC without Adjuvant Therapy



[Print this Page](#)

Presentation Abstract

Abstract
Number: 5556

Presentation
Title: **Identification of signaling pathways active in small cell lung cancer (SCLC) compared to non-small cell lung cancer (NSCLC) by proteomic profiling**

Presentation
Time: Wednesday, Apr 21, 2010, 8:00 AM -11:00 AM

Location: Exhibit Hall A-C, Poster Section 28

Poster
Section: 28

Poster Board
Number: 3

Author
Block: Lauren Averett Byers¹, Jing Wang¹, John Yordy¹, You-Hong Fan¹, Uma Giri¹, Li Shen¹, Ignacio Wistuba¹, Luc Girard², Kevin Coombes¹, John Weinstein¹, John Minna², John Heymach¹. ¹M.D. Anderson Cancer Center, Houston, TX; ²University of Texas Southwestern, Houston, TX


Abstract
Body: **Introduction:** While there have been multiple studies identifying differences in phenotypes of small cell lung cancer (SCLC) compared to non-small cell lung cancer (NSCLC), identification of key signaling pathways that may differ and may represent new therapeutic targets for SCLC need to be specified. To this end, we have compared expression of a large number of proteins and genes involved in key cell regulatory pathways of SCLC with those in NSCLC using reverse phase protein array (RPPA) and genome-wide mRNA profiling.

Methods: 34 SCLC and 74 NSCLC cell lines that have been previously characterized for oncogenotype mutations, copy number changes, and mRNA expression profiles (using Affymetrix U133A, U133B, U133 Plus 2.0, and Illumina arrays) were grown under 10% serum, serum-starved, and serum-stimulated conditions. Cell lysates from each media condition were harvested and 176 total and phosphoproteins representing multiple signaling pathways were quantified by RPPA. Differences in protein and gene expression levels between SCLC and NSCLC lines were assessed using unsupervised hierarchical clustering, t-test, and ANOVA.

Results: Unsupervised hierarchical clustering separated SCLC from NSCLC cell lines based on their distinct protein expression patterns under all three media conditions. Two high-grade neuroendocrine NSCLC cell lines clustered with the SCLC cell lines, reflecting their SCLC-like protein expression. Using a false discovery rate of <1% to identify protein markers most differentially expressed between SCLC and NSCLC, we observed high levels of proteins in DNA repair pathways (ex., PARP1, XRCC1, ATM, Chk1, 53BP1) and in apoptosis pathways (ex., Bcl-2, BIM, BAX). In contrast, MAPK pathway activity was downregulated, as reflected by low levels of phospho-ERK1/2. Among SCLC lines, an inverse correlation was observed between cKit and cMyc protein levels such that one group overexpressed cKit but had relatively lower cMyc levels, while the other group (enriched for cMyc amplification) had high cMyc and low cKit. Gene expression analysis showed similar results, with good correlation between non-phosphorylated proteins and mRNA levels among the pathways differentially expressed in SCLC.

Conclusions: Proteomic profiling identified increased DNA repair and apoptosis pathway activity and decreased MAPK pathway activity in SCLC. Among the proteins upregulated in SCLC, several have potential as therapeutic targets, such as PARP1. Although cKit over expression and cMyc amplification are both common in SCLC, we showed that they were mutually exclusive of each other at a protein expression level, reflecting the heterogeneity of SCLC and suggesting a rationale for testing targeted therapies specifically in subsets of patients whose tumors overexpress the targeted marker. Pathway activation will be further investigated in SCLC patient tumors.

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106



INTERNATIONAL ASSOCIATION FOR THE STUDY OF LUNG CANCER

IASLC

13th World Conference on Lung Cancer JULY 31 - AUGUST 4, 2009 MOSCONE WEST SAN FRANCISCO, USA

[Programme overview](#)
[Authors](#)
[Moderators](#)


[« Back](#)

» Select day

» Please select a type

» Select topic


Tuesday, August 4, 2009



BOOK MARK

D10

Anti-Angiogenesis and Anti-Coagulants
Moscone West 2001 - 2005, Level 2: 12:30 - 14:00
Type: Oral
 Moderation: P.A. Janne, H.(. West;



BOOK MARK

D10.1 - Stromal HGF and VEGFR-1 are associated with acquired resistance to VEGFR tyrosine kinase inhibitors in non-small cell lung cancer (NSCLC)
Oral
Novel Therapeutics and Modalities
T. Cascone, M. Herynk, D. Du, H. Kadara, E. Hanrahan, M. Nilsson, H.Y. Lin, J..J. Lee, Y.-Y. Park, J.-S. Lee, J.V. Heymach; Houston, TX/US

Body

Background: Tyrosine kinase inhbitors (TKIs) targeting vascular endothelial grow factor/receptor (VEGF/VEGFR) pathway, such as cediranib (AZD2171) and the dual VEGF 2/EGFR inhibitor vandetanib (ZD6474) have demonstrated clinical benefit in NSCLC and oth solid tumors. Unfortunately, while a subset of patients initially responds to these agent therapeutic resistance inevitably emerges. The mechanisms underlying resistance are not w understood, but may include incomplete inhibition of VEGFR or EGFR activation, bypass these pathways through redundant expression of proangiogenic molecules, a switch in strom cell dependency from VEGFR activation to alternative signaling pathways or other mechanism Therefore, there is a critical need for tumoral and stromal molecular markers able to pred resistance to TKIs such as vandetanib and cediranib in NSCLC.Methods: To generate NSCL xenograft models with acquired resistance to cediranib and vandetanib, A549 and H197 human NSCLC cells were injected in athymic nude mice. Tumors were considered resistant treatments when their volumes tripled compared to the starting volume and were th sacrificed. A murine specific expression array (mouseWG-6 v2 Expression BeadChip, ILLUMIN was used to evaluate changes in the stromal microenvironment to develop gene signatur indicative of therapeutic resistance in vandetanib and cediranib-resistant H1975 tumo compared to sensitive tumors (2 weeks of treatment) and vehicle treated contro Differentially expressed genes between the different treatment groups were selected based on $p < 0.005$ of the univariate t-test and at least a 1.5 fold change in expression Results: H197

[Print this Page](#)

Presentation Abstract

Abstract
Number: 376

Presentation Title: Increased HGF is associated with resistance to VEGFR tyrosine kinase inhibitors (TKIs) in non-small cell lung cancer (NSCLC)

Presentation Time: Sunday, Apr 18, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 13

Poster
Section: 13

Poster
Board
Number: 13

Author Block: Tina Cascone¹, Matthew H. Herynk¹, Li Xu¹, Humam N. Kadara¹, Emer Hanrahan¹, You-Hong Fan¹, Babita Siagal¹, Yun-Yong Park¹, Ju-Seog Lee¹, Robert R. Langley¹, Juliane M. Jurgensmeier², Anderson Ryan², John V. Heymach¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²AstraZeneca, Macclesfield, United Kingdom

Abstract Body: Emergence of therapeutic resistance to angiogenesis inhibitors, the mechanisms of which are poorly understood, remains a major obstacle in treatment of NSCLC patients. Previously we reported that mechanisms governing resistance to anti-angiogenic therapy may involve both tumor and stromal cells in the tumor microenvironment. In this study we investigated potential mechanisms of resistance to the multi-tyrosine kinase inhibitors cediranib (AZD2171, Recentin[®]) and vandetanib (ZD6474, Zactima[®]) using NSCLC xenografts treated either for 2 weeks (sensitive tumors) or until resistance occurred. Quantification of TUNEL⁺ staining using laser scanning cytometry (LSC) showed increased apoptosis in H1975 xenografts sensitive to cediranib ($p < 0.01$) and vandetanib ($p < 0.05$) when compared with controls, whereas no changes were noticed at time of resistance. Microvessel density (MVD) was significantly increased in resistant H1975 xenografts compared with controls ($p < 0.05$) and sensitive tumors ($p < 0.01$), whereas in A549 model, vandetanib-resistance was associated with an angiogenic independent phenotype. To investigate stromal mechanisms of Vascular Endothelial Growth Factor Receptor (VEGFR) TKI resistance, we characterized the stromal angiogenic gene expression profiles of H1975 sensitive and resistant tumors using a mouse-specific gene expression array (mouseWG-6 v2 Expression BeadChip, Illumina[®]). Differentially modulated genes were selected based on a $p < 0.005$ of the univariate t-test and at least a 1.5 fold-change in expression and cross-referenced to defined list(s) of angiogenesis-related genes. Stromal *Hgf* (hepatocyte growth factor) was up-regulated in VEGFR TKI-resistant xenografts compared to sensitive tumors. *Hgf* up-regulation was confirmed at the protein level using immunofluorescent staining and confocal microscopy. HGF protein levels were strongly decreased after 2 weeks of treatment with cediranib and vandetanib ($p < 0.01$), whereas a significant increase in HGF was observed in resistant xenografts ($p < 0.01$). To assess whether HGF upregulation contributes to tumor resistance to TKIs, we implanted HGF-overexpressing and vector control HCC827 NSCLC cells into nude mice. In HCC827-vector control xenografts, cediranib inhibited tumor growth by 93%, whereas a 60% of growth inhibition was observed in HGF-overexpressing tumors. These data agree favorably with our previous analysis of clinical specimens from patients with stage IIIB/IV NSCLC that identified HGF as a predictive marker of resistance to vandetanib treatment alone when compared to chemotherapy or the combination of chemotherapy and vandetanib ($p = 0.033$). Our results suggest that HGF up-regulation may resistance to VEGFR pathway inhibition and that the HGF/MET axis may represent a crucial target for NSCLCs that are resistant to anti-angiogenic therapy.

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Profiling in pharmacologically re-expressed microRNAs by 5-azacytidine and SAHA identified metastases associated miR-148b in malignant pleural mesothelioma cell line

Corvalan A, Suraokar M, Xiaoyan Zou, Chow C-W, Gazdar A, Moran C, Raso G, Mehran R, Tsao A, Wistuba I.

Background. MicroRNAs (miRNAs) have emerged as key players in human carcinogenesis. Recently it has been shown that some miRNAs can be epigenetic up-regulated by aberrant hypermethylation in human cancer. Malignant pleural mesothelioma (MPM) is a highly malignant neoplasm with different histological subtypes. To explore the role of epigenetic mediated up-regulation of miRNAs in MPM we performed pharmacological unmasking of miRNA expression in cell lines.

Methods Five mesothelioma cell lines, including one normal mesothelial (Met5A) and five MPMs (epithelioid H2452, biphasic H211 and unclassified H28 and H2052) were treated with the demethylating agent 5-aza-cytidine (5-Aza;1 uM) and SAHA (2.5 uM) for 96 hrs. After RNA extraction (Trizol) miRNA profiling was performed by Agilent human microRNA kit v2.

Results Total miRNA up-regulated (two-fold) after the treatment were 299 (51%) in normal mesothelial Met5A cell line, and lower in the malignant cell lines: 171 (29%) in H2452, 79 (13.5%) in 211H, 55 (9.4%) in H28, and 56 (9.6%) in H2052. Interestingly 167 (55.9%) miRNAs were exclusively up-regulated in Met5A, 56 (32.7%) in H2452, 21 (26.6%) in 211H, 16 (29.1%), in H28, and 18 (32.1%) in H2052. Among all unique miRNA, only 17 (let-7b, let-7c, let-7f-2, miR-302c, miR-328, miR-510, miR-125b-1, miR-16-1, miR-223, miR-302b, miR-383, miR-551b, miR-922, miR-148a, miR-18b, miR-302d, miR-326) have been previously associated to human carcinogenesis. Interestingly one of these miRNA (miR-148a) has been associated with microRNA metastasis signature. **Discussion** The number of total and unique miRNA upregulated after 5-Aza and SAHA was lower in malignant mesothelioma cell lines vs normal Met5A cell line. Up-regulation of unique miRNAs was found associated with specific subtypes of MPM. The identification of metastasis associated miR-148a suggests a potential biomarker for metastasis in this highly malignant neoplasm. Further research, including tissue specimens, will be necessary to validate these results. **Grant support** PROSPECT DoD W81XWH-07-1-0306.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 1624

Presentation Title: Identification and characterization of phospho-ORM-1 as a novel nicotinic acetylcholine receptor (nAChR)-associated protein and a potential serum marker for lung cancer detection

Presentation Time: Monday, Apr 19, 2010, 9:00 AM -12:00 PM

Location: Exhibit Hall A-C, Poster Section 25

Poster
Section: 25

Poster
Board
Number: 24

Author Block: Gita Jayachandran, Xifeng Wu, Kevin Coombes, Ignacio Wistuba, Jack Roth, Lin Ji. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract Body: Phosphorylation is a key regulating switch that plays a critical role in signaling pathways involved in oncogenesis. A high-throughput analysis of phosphoproteins could provide a useful tool for analysis of biological functions and signaling pathways associated with these molecular events. In this study, we developed an innovative functional proteomics platform using ProteinChip array-based SELDI-MS for high-throughput profiling and identification of phosphopeptides in human serum to identify specific phosphopeptides/phosphoproteins associated with human lung cancer. We performed phosphopeptide profiling on serum samples from human normal and lung cancer patients with varying stages and smoking histories. We used phospho-tyrosine antibody-conjugated super-paramagnetic beads to capture phosphopeptides generated in trypsin-digested serum samples. The captured phospho-tyrosine peptides (pYs) were separated on a hybrid magnetic plate using a biological sample preparation robot. The affinity-enriched pYs were then randomly loaded onto ProteinChips and analyzed by SELDI-TOF Mass Spectrometry. We used wavelets and the mean spectrum for peak detection and detected more than 600 pYP peaks spanning M/Z range from 50 to 5500 Dalton. For each peak, we recorded the p-value from an F-test and modeled the set of p-values using a beta-uniform mixture model to estimate the false discovery rate (FDR). We identified 39 pYP peaks with fold changes in intensity detected on SELDI-MS profiles to be significantly (at FDR = 10%) differentially expressed between the normal and lung cancer serum samples. The phosphopeptides detected on SELDI-MS spectra were further identified using a protein chip array-interfaced qSTAR-MS/MS. One of phospho-tyrosine containing peptides was identified as an Alpha-1-acid glycoprotein 1 precursor (A1AG1) or ORM-1 (Orosomucoid). The ORM-1 pYP showed a M/Z peak at 1752.3 Da and was significantly upregulated in lung cancer serum samples with more than 10-fold increase ($P = 0.0024$) in mass peak intensity. Computer-aided structural and functional analysis predicted the potential association of ORM-1 to the nicotinic acetyl choline receptor (nAChRs). We further validated phospho-ORM-1 protein expression in another set of lung cancer and control serum samples by Quantitative ELISA and confirmed the significantly upregulated expression of serum phospho-ORM-1 in lung cancer patients with ever-smoking history. We also identified protein interactions between the ORM-1 and subunits of nAChRs in lung cancer cell lines by immunoprecipitation and immune-blotting analysis. Our results suggest the role of the Phospho-ORM-1 peptide as a novel nAChR-associated protein in lung cancer pathogenesis and smoking-associated carcinogenesis.

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Enriched Tumor Expression of Folate Transporters Correlates With Adenocarcinoma Histology Type, Female Gender and Presence of *EGFR* Mutation in Non-Small Cell Lung Carcinoma

Author Block: *Maria Ines Nunez, Carmen Behrens, Denise M. Woods, Heather Lin, Milind Suraokar, Luc Girard, John Minna, Jack Lee, W Hofstetter, Wilbur Franklin, Cesar A. Moran, Waun K. Hong, David J. Stewart, Ignacio I. Wistuba.* Pathology Department - UT M.D. Anderson Cancer Center, Houston, TX, Biostatistics Department - UT M.D. Anderson Cancer Center, Houston, TX, Hamon Center - UT Southwestern Medical Center, Dallas, TX, Thoracic Surgery Department - UT M.D. Anderson Cancer Center, Houston, TX, University of Colorado Cancer center, Denver, CO, Thoracic/Head and Neck Medical Oncology Department - UT M.D. Anderson Cancer Center, Houston, TX

Background: Membrane bound folate receptor alpha (FOLR1) and transmembrane Reduced folate carrier 1 (RFC1) regulate uptake of folate as well as folate linked conjugates inside the cell. FOLR1 and RFC1 are over expressed in epithelial primary and metastatic tumors and are promising therapeutic targets and tumor biomarkers. Due to limited information on the expression of these receptors in non-small cell lung carcinoma (NSCLC) we studied the protein immunohistochemical (IHC) expression of these receptors in a large set of tumors and correlate our findings with patients' clinicopathologic features.

Methods: IHC protein expression of FOLR1, RFC1, was examined in 320 surgically resected NSCLCs placed in tissue microarrays, including 202 adenocarcinomas and 110 squamous carcinomas, and correlated with patients' clinico-pathological characteristics. A semiquantitative IHC score was obtained assessing intensity of immunostaining and percentage of positive tumor cells.

Results: The pattern of IHC expression varied in malignant cells, with FOLR1 and RFC1 expressed in the membrane and cytoplasm. In all cases expression in tumor cells was higher than in non-malignant lung epithelial cells. Tumor stroma IHC expression was frequently detected, especially in endothelial cells, lymphocytes, macrophages and fibroblasts. Adenocarcinomas showed significantly higher expression compared with squamous cell carcinoma for membrane ($P<0.001$) and cytoplasmic ($P<0.001$) FOLR1. Female NSCLC patients had significantly higher expression of membrane and cytoplasmic FOLR1 ($P=0.01$) compared with male patients. Ever smoker patients demonstrated significantly lower expression of membrane ($P<0.001$) and cytoplasmic FOLR1 ($P<0.002$), and higher expression of membrane RFC1 ($P=0.01$), compared with never smokers. In adenocarcinomas, the presence of *EGFR* mutations correlated with higher expression of membrane FOLR1 ($P<0.002$). Finally, squamous carcinomas showed higher positive endothelial cell expression of FOLR1 ($P=0.00001$) than adenocarcinomas.

Conclusion: 1. FOLR1 and RFC1 membrane transporters proteins are over expressed in NSCLC compared to normal lung epithelium; 2. significant differences were found between adenocarcinomas and squamous lung cancer in both tumor cells and the tumor microenvironment; 3. differences were found in tumors of males and females, between tumors from never and ever smokers, and tumors with *EGFR* or *KRAS* mutations. The different patterns of transporter expression may explain the superior response of NSCLC patients with adenocarcinoma histology to pemetrexed.

Supported by grants US DoD W81XWH-07-1-0306, and UT-Lung SPORE P50CA70907.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 1981

Presentation Title: MYC downregulation and chemoresistance in non-small cell lung cancer (NSCLC): Evidence from the Biomarker-Based Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE) program

Presentation Time: Monday, Apr 19, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 1

Poster
Section: 1

Poster
Board
Number: 4

Author Block: Pierre Saintigny¹, Lauren A. Byers¹, Li Zhang¹, John S. Yordy¹, Xi M. Tang¹, Luc Girard², Wenhua Lang¹, You H. Fan¹, Lin Ji¹, Jack J. Lee¹, Edward S. Kim¹, Waun K. Hong¹, Scott M. Lippman¹, Roy S. Herbst¹, John Minna², Ignacio I. Wistuba¹, John V. Heymach¹, Li Mao¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²UT Southwestern Medical Center, Dallas, TX

Abstract Body: Background: Primary or acquired resistance to platinum-based therapy remains an important issue in the treatment of unresectable non-small cell lung cancer (NSCLC). Our objective was to identify factors of resistance to platinum-based therapy. Experimental procedures: Gene expression profiling from the BATTLE program (pretreated resistant tumors stage III/IV, N=32) and a control group from 2 independent publicly available datasets (never treated tumors stage III/IV, N=45) were compared. Gene expression profiling was generated using the same platform (U133 Plus 2.0 Array). Pathway and gene set analyses were used to define networks and pathways associated with resistance. For validation, we used: i-an independent set of 38 never treated NSCLC stage III/IV, ii-a set of 53 NSCLC cell lines tested for cisplatin sensitivity with proteomic profiling generated by Reverse Phase Protein Array (RPPA) technology using 177 well-characterized antibodies, iii-the comparison of two pairs of NSCLC cell lines (H1437, H460) that were made resistant by iterative exposure to cisplatin, and iv-transfection experiments. Results: A total of 3,963 probesets were found to be differentially expressed between BATTLE samples and the control group with a p-value < 0.001 (two-tailed t-test). DNA repair gene sets were upregulated in BATTLE samples compared to never treated tumors. Network analysis found that MYC as well as many of its downstream regulated genes were significantly downregulated in BATTLE samples. In particular, a high proportion of MYC target genes associated with apoptosis were downregulated. Using real time PCR, MYC gene expression was significantly lower in 15 BATTLE samples compared with an independent set of 38 never treated stage III/IV NSCLC (p-value=0.01). The proteomic profiling of 53 NSCLC cell lines showed that MYC expression was one of the top proteins associated with sensitivity to cisplatin. An inverse correlation was observed between sensitivity to cisplatin (IC50) and MYC protein expression evaluated by RPPA (total MYC: r=-0.41, p-value=0.003; phosphorylated MYC: r=-0.30, p-value=0.03). MYC gene expression by real time PCR was lower in resistant H1437 and H460 cell lines compared to parental cell lines. Finally, transfection of H226 cell lines with a vector expressing MYC improved sensitivity to cisplatin compared to the control. Conclusion: This is the first gene expression profiling analysis of NSCLC samples pretreated and resistant to chemotherapy included in a prospective clinical trial. MYC down regulation may play an important role in NSCLC resistance to platinum-based chemotherapy. Mechanisms of MYC down regulation should be further explored.

American Association for Cancer Research
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

[Print this Page](#)

Presentation Abstract

Abstract Number: 4648

Presentation Title: High membrane carbonic anhydrase IX (CAIX) expression predicts relapse of resected stage I-II non-small cell lung cancer (NSCLC)

Presentation Time: Tuesday, Apr 20, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 31

Poster Section: 31

Poster Board Number: 11

Author Block: David J. Stewart, Maria I. Nunez, Carmen Behrens, Heather Lin, J. Jack Lee, Jack Roth, John Victor Heymach, Diane D. Liu, Ivan Ignacio Wistuba. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract Body: **Background:** Stage predicts outcome in NSCLC but some stage I-II patients relapse. Additional prognostic factors are needed. **Methods:** From our NSCLC tumor bank we selected tissues of 230 stage I-II patients who had not received adjuvant chemotherapy. We assessed by immunohistochemistry (IHC) selected factors related to cell growth rate and regulation, hypoxia, transporters and DNA repair. IHC scores (0-300) were calculated by multiplying stain intensity (0-3) by % tumor cells stained. We assessed nuclear (N) p53, p21WAF1/CIP1, Ki67; cytoplasmic (C) COX2, DcR2; N and C CTR1, DNMT1, HIF1a, Rb, pRb, SHARP2, Survivin, VEGF, p14ARF, p16INK4, ERCC1; C and membrane (M) CAIX, TGF-beta. Definitions were: Time to Relapse (TTR): time from surgery to relapse or last follow up (LFU), with censoring at LFU or death or diagnosis of a metastatic 2nd primary, if clinically relapse-free from initial NSCLC; Overall Survival (OS): time from surgery to death, with censoring at LFU if alive at that time. **Results:** Exponential decay nonlinear regression analysis [EDNRA] of TTR curves suggested that 70% of patients were cured, with a TTR half-life = 20 months for those relapsing. By EDNRA, OS was uniphasic with half-life = 89 months and no indication of a survival inflection point differentiating those dying from NSCLC vs other causes. In multivariate Cox models, factors correlating with high recurrence risk were M CAIX (any vs none, hazard ratio [HR] 2.08, p=0.02) and node stage (N1 vs N0, HR 2.59, p=0.002). M CAIX (HR 1.92, p=0.05) and node stage (HR 2.54, p=0.003) remain in the model if tumor diameter (TD) is forced into the model (TD HR=1.068, p=0.36). M CAIX correlated (p<0.05) directly with TD, squamous vs adenocarcinoma, C CAIX, C SHARP2, N DNMT1, N pRb and N Ki67 and inversely with C and N CTR1, C and N p16INK4, C DNMT1, C HIF1a, C VEGF, C Rb and N p14ARF. In multivariate Cox models, short OS correlated with age (HR 1.05, p<0.0001), smoking history (HR 2.56, p=0.01), node stage (HR 1.67, p<0.05), TD (HR 1.16, p<0.03), high C CAIX (HR 1.004, p=0.009), low C Rb (HR 0.993, p<0.0001) and high C pRB (HR 1.005, p=0.03). In univariate analyses, factors correlating (p<0.05) with high node stage were high TD and N Rb, and low C DNMT1 and C Rb, with trends (p<0.10) to correlations with high M CAIX, N Survivin, N p53 and N pRb and with low C CTR1, C HIF1a and N p16INK4. Factors correlating (p<0.05) with high TD included high node stage and M CAIX, low C VEGF and other factors. **Conclusion:** CAIX deserves further attention as a prognostic factor and therapy target, and correlated more strongly with TTR than did TD or other markers, including the hypoxia markers VEGF and HIF1a. The lack of inflection points on OS EDNRA curves suggests that some of the factors may predict comorbidity and not tumor biology. Hence, TTR as defined here may be the preferred endpoint to assess tumor biology. Support: DoD grant # W81XWH-07-1-0306 and UT-Lung SPORE P50CA070907.

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

[Print this Page](#)

Presentation Abstract

Abstract
Number: 2172

Presentation
Title: Deregulation of the mitotic spindle assembly checkpoint pathway in malignant pleural mesothelioma (MPM) tumors revealed by gene expression profiling

Presentation
Time: Monday, Apr 19, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 8

Poster
Section: 8

Poster
Board
Number: 13

Author
Block: Milind B. Suraokar, Maria Nunez, Chi-wan B. Chow, Gabriela Mendoza, Lixia Diao, Cesar Moran, Gabriela Raso, Reza Mehran, Alejandro Corvalan, Kevin Coombes, Anne Tsao, Ignacio Wistuba. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract
Body: MPM is a lethal neoplasm exhibiting low median survival of patients and lacks effective therapeutic options. There is an urgent need to understand the underlying pathobiology and discover novel therapeutic targets. We undertook a messenger RNA (mRNA) expression profiling strategy to determine the pathways and biomarkers significantly altered in MPM tumors. We isolated total RNA from 55 MPM tumors with 38 paired controls representing 39 epitheloid, 8 biphasic and 6 sarcomatoid cases. The paired controls were adjacent non-tumor tissue samples and histopathological analysis revealed a tumor content of greater than 70 % in most of these tumors cases. The RNA was labeled and hybridized onto Affymetrix U133 Plus 2.0 chips to ascertain the global expression profile in these tumors. Bioinformatic analysis of the microarray data using a two sample t-test was applied on a probe-by-probe basis followed by Beta-uniform Mixture for multiple comparisons. Finally paired t-test was applied to determine the differences between tumor vs normal samples. About ~955 highly significant probesets representing ~ 670 genes, at a FDR of e-09, were obtained and subjected to pathway analysis using MetaCore software suite (GeneGo, Inc.). The most significantly altered pathway in MPM tumors was the Mitotic Spindle Assembly Checkpoint (MSAC) pathway due to up-regulation of at least 15 genes including a ~3.4 fold increase in Aurora kinase A, which is currently being explored in other cancers as a potential therapeutic target. The other genes belonging to this pathway, also up-regulated in tumors, include Mad2L1 and BUBR1 which together regulate cell division cycle 20 (Cdc20) protein, an essential cofactor needed by the Anaphase Promoting Complex to initiate the anaphase of cell cycle. Interestingly we also discovered that some of the MSAC pathway genes show a histotype-specific graded expression pattern with higher levels in sarcomatoid tumors compared to biphasic tumors and with lowest expression levels seen in epitheloid tumors. Additionally survivin, the product of which localizes to the mitotic spindle and negatively regulates apoptosis by inhibiting caspase activation, was expressed more than 2 fold in MPM tumors compared to normal samples. The microarray data also revealed other pathways significantly upregulated in MPM tumors including the Wnt and Cell-Adhesion signaling pathways. We are currently validating the expression microarray data using quantitative Polymerase Chain Reaction (PCR) platform with respect to key components of MSAC pathway. This will be followed by proteomic analysis of these components on tumor lysates to confirm the alterations in their expression profiles. Supported by Grants: DOD PROSPECT W81XWH-07-1-030602, Fleming Foundation for Mesothelioma Research, Aileen Dillon and Lee Bourg Mesothelioma Endowment and NIH K12 CA088084 08 award.

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Evaluation of Src Tyr419 as a predictive biomarker in a neoadjuvant trial using dasatinib in resectable malignant pleural mesothelioma.

Sub-category:

Local-Regional Therapy

Category:

Lung Cancer - Local-Regional and Adjuvant Therapy

Meeting:

2010 ASCO Annual Meeting

Session Type and Session Title:

General Poster Session, Lung Cancer - Local-Regional and Adjuvant Therapy

Abstract No:

7042

Citation:

J Clin Oncol 28:15s, 2010 (suppl; abstr 7042)

Author(s):

A. S. Tsao, I. I. Wistuba, R. J. Mehran, J. Gil, M. Nunez, J. Lee, H. Lin, S. M. Lippman, W. K. Hong; University of Texas M. D. Anderson Cancer Center, Houston, TX

Abstract:

Background: Preclinical studies in malignant pleural mesothelioma (MPM) have shown that dasatinib, a multitargeted Src kinase/PDGFR inhibitor, has antitumor activity. We designed a novel biomarker-based neoadjuvant trial using dasatinib in resectable MPM. **Methods:** Untreated MPM patients undergo extended surgical staging (ESS) with multiple biopsies along the future surgical incision line to account for tumor heterogeneity. If deemed a surgical candidate for either P/D or EPP, patients receive 4 weeks of oral dasatinib (70 mg BID) followed by P/D or EPP. If a radiographic response is seen, an additional 2 years of dasatinib maintenance after adjuvant radiotherapy and chemotherapy is given. Serum/blood/platelets/pleural effusion are collected for exploratory analysis of peripheral surrogate biomarkers. Primary endpoint is biomarker modulation of Src Tyr419 in tumor tissue. Secondary endpoints: response, survival, safety/toxicity, and peripheral biomarker modulation. **Results:** 15 patients enrolled on the trial (4/08-1/10) have successfully completed ESS, neoadjuvant dasatinib, and P/D (n=10) or EPP (n=5). 13 epithelioid and 2 biphasic histology, 12 men: 3 women. The dasatinib main side effects were grade 1-2: anemia, nausea, vomiting, anorexia, fatigue, and anxiety. Grade 3 toxicities included fluid retention, infection (pneumonia), and hypoxia. There are no grade 4-5 toxicities. Post-surgical grade 3 toxicity included anemia, arrhythmia, HTN, and pleural effusion; 1 grade 4 episode of hyperglycemia. After 4 weeks of neoadjuvant dasatinib therapy, there was 1 PD, 12 SD and 2 minor responses. In the initial analysis of Src Tyr419 in 13 patients, higher baseline levels of p-Src Tyr419 predicted for an improved PFS with dasatinib therapy (p=0.008). Also, patients who had significant modulation of p-Src Tyr419 after dasatinib therapy had improved PFS (p=0.008). **Conclusions:** There is preliminary evidence that a subgroup of MPM patients gain clinical benefit from dasatinib therapy and that baseline p-Src Tyr419 levels in MPM tumor tissue may be predictive of PFS. This is the first targeted therapy neoadjuvant trial to potentially identify a predictive biomarker in MPM.

Abstract Disclosures**Faculty & Discussant Disclosures****Annual Meeting Planning Committee Disclosures**

Abstracts that were granted an exception in accordance with ASCO's Conflict of Interest Policy and are designated with a caret symbol (^) here and in the print version.

► **Associated Presentation(s):**

1. Evaluation of Src Tyr419 as a predictive biomarker in a neoadjuvant trial using dasatinib in resectable malignant pleural mesothelioma.

Meeting: [2010 ASCO Annual Meeting](#)

Presenter: [Anne S. Tsao](#)

Session: [Lung Cancer - Local-Regional and Adjuvant Therapy \(General Poster Session\)](#)

► **Other Abstracts in this Sub-Category:**

1. A multi-institutional feasibility study of induction pemetrexed (Pem) plus cisplatin (Cis) followed by extrapleural pneumonectomy (EPP) and postoperative hemithoracic radiation (RT) for malignant pleural mesothelioma (MPM): Japan Mesothelioma Research Center (JMRC) 001001 trial.

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: TPS285 First Author: [F. Tanaka](#)

Category: [Lung Cancer - Local-Regional and Adjuvant Therapy - Local-Regional Therapy](#)

2. A dose-escalation study of vorinostat in combination with radiotherapy for patients with non-small cell lung cancer.

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: TPS286 First Author: [R. H. Decker](#)

Category: [Lung Cancer - Local-Regional and Adjuvant Therapy - Local-Regional Therapy](#)

3. Trial of poor performance status patients (ToPPS): A randomized phase II trial of pemetrexed versus pemetrexed/bevacizumab versus pemetrexed/carboplatin/bevacizumab in patients with stage IIb/IV non-small cell lung cancer and ECOG performance status 2.

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: TPS287 First Author: [D. R. Spigel](#)

Category: [Lung Cancer - Local-Regional and Adjuvant Therapy - Local-Regional Therapy](#)

[More...](#)

► **Abstracts by A. S. Tsao:**

1. Evaluation of Src Tyr419 as a predictive biomarker in a neoadjuvant trial using dasatinib in resectable malignant pleural mesothelioma.

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: 7042 First Author: [A. S. Tsao](#)

Category: [Lung Cancer - Local-Regional and Adjuvant Therapy - Local-Regional Therapy](#)

2. Human papillomavirus (HPV) transmission from oropharyngeal cancer patients to sexual partners.

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: 5527 First Author: [V. Papadimitrakopoulou](#)
Category: [Head and Neck Cancer](#)

3. [Sorafenib treatment efficacy and KRAS biomarker status in the Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination \(BATTLE\) trial.](#)

Meeting: [2010 ASCO Annual Meeting](#) Abstract No: 7609 First Author: [R. S. Herbst](#)
Category: [Lung Cancer - Metastatic - Metastatic](#)
[More...](#)

► **Presentations by A. S. Tsao:**

1. Evaluation of Src Tyr419 as a predictive biomarker in a neoadjuvant trial using dasatinib in resectable malignant pleural mesothelioma.

Meeting: [2010 ASCO Annual Meeting](#)
Presenter: [Anne S. Tsao](#)
Session: [Lung Cancer - Local-Regional and Adjuvant Therapy](#) (General Poster Session)

2. Activated Src kinase is expressed in malignant pleural mesothelioma tumors; dasatinib inhibition leads to cytotoxicity, cell cycle inhibition, and prevention of invasion and migration.

Meeting: [2007 ASCO Annual Meeting](#)
Presenter: [Anne S Tsao, MD](#)
Session: [Lung Cancer](#) (General Poster Session)

3. Phase I/II trial of bevacizumab plus erlotinib for patients with recurrent non-small cell lung cancer: Correlation of treatment response with mutations of the EGFR tyrosine kinase gene

Meeting: [2005 ASCO Annual Meeting](#)
Presenter: [Anne S Tsao, MD](#)
Session: [Lung Cancer](#) (General Poster Session)

[More...](#)

► **Educational Book Manuscripts by A. S. Tsao:**

No items found.

[Print this Page](#)

Presentation Abstract

Abstract
Number:

4119

Presentation
Title:

Cytoplasmic location of CXCR4 is correlated to loss of EMT marker and activation of downstream signaling pathway in non-small cell lung cancer

Presentation
Time:

Tuesday, Apr 20, 2010, 2:00 PM - 5:00 PM

Location:

Exhibit Hall A-C, Poster Section 10

Poster
Section:

10

Poster
Board
Number:

23

Author
Block:

Yuxiang Wang¹, Donghai Huang¹, Hongpeng Fu², Dongsheng Wang¹, Lydia Koenig¹, Fadlo R. Khuri¹, Dong M. Shin¹, Zhuo (Georgia) Chen¹. ¹Emory University, Atlanta, GA; ²China Health Economic Institute, Beijing, China

Abstract
Body:

Introduction: Although epidermal growth factor receptor (EGFR) targeted therapy initially proved to be effective in some lung cancer patients, emerging resistance to EGFR- tyrosine kinase inhibitor (TKI) became an urgent issue to solve. Some research has shown that cancer cells undergoing epithelial mesenchymal transition (EMT) acquired resistance to EGFR-TKI, though the underlying mechanism remains unknown. Our preliminary study in non-small cell lung cancer (NSCLC) cell lines showed that G-protein coupled chemokine receptor CXCR4 could activate both p-AKT and p-ERK pathways, which have been reported to induce EMT. In this study, we hypothesize that CXCR4 may be an alternative route for cancer cells to bypass EGFR to activate the downstream pathway under an EMT phenotype. **Methods:** Immunohistochemistry (IHC) was used to detect the expressions of CXCR4, p-AKT, p-ERK, and E-cadherin in 94 clinical NSCLC samples and a weighted index (WI = % expression x intensity score) was used to quantify the expression level of these markers. Western blot was used to detect the expression of CXCR4 and activation of its downstream pathways in NSCLC cell lines. **Results:** CXCR4 was detected in 98.9% (93/94) of lung cancer samples, however, of greater interest were the 3 distinctive expression patterns for CXCR4 including cytoplasmic, nuclear, and cytoplasmic/nuclear expressions. Since it seems unreasonable to see a transmembrane receptor located in nuclei, we confirmed nuclear localization of CXCR4 by western blot using nuclear extractions from NSCLC cell lines. Furthermore, we evaluated the correlation of different locations of CXCR4 with p-AKT, p-ERK, and E-cadherin by ANOVA statistical analysis. The result showed the expression level of p-AKT and p-ERK are both significantly higher ($p=0.004$ and $p=0.014$, respectively) in CXCR4 cytoplasmic and cytoplasmic/nuclear localization groups than those in CXCR4 nuclear localization group. The expression level of E-cadherin in the CXCR4 cytoplasmic and cytoplasmic/nuclear localization groups was also significantly lower ($p=0.021$) than that in CXCR4 nuclear staining group. **Conclusions:** Our data suggested that cytoplasmic CXCR4 may contribute to activation of AKT and ERK signaling pathways under EMT, while nuclear localization of CXCR4 may be an inactive form. The functional cytoplasmic CXCR4 may contribute to EGFR-TKI resistance by offering an alternative pathway for tumors to maintain progression. (Supported by grants from NIH R21 CA125062, DOD W81XWH-07-1-0306 Project 5 to ZGC, and GCC Distinguished Scholar Award to ZGC and DMS).

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

[Print this Page](#)

Presentation Abstract

Abstract
Number: 374

Presentation
Title: ***VEGFR2/KDR* polymorphisms and gene copy gain with outcome in non-small cell lung cancer**

Presentation
Time: Sunday, Apr 18, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 13

Poster
Section: 13

Poster
Board
Number: 11

Author
Block: Fei Yang¹, Ximing Tang¹, Alejandro Corvalan¹, Carmen Behrens¹, Heather Lin¹, Maria Gabriela Raso¹, J. Jack Lee¹, John D. Minna², Ignacio Ivan Wistuba¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²UT Southwestern Medical Ctr., Dallas, TX

Abstract
Body: Background. The VEGF receptor 2 (VEGFR2) is the predominant mediator of VEGF-stimulated endothelial cell function. Recently, *VEGFR2/KDR* copy number gain and mutation, and *VEGF* copy number gain, have been described in lung adenocarcinoma tumor specimens. To better characterize the molecular changes of these two genes in non-small cell lung carcinoma (NSCLC), we investigated their abnormalities in NSCLC tumor tissue specimens and correlated with patients clinico-pathological characteristics.
Methods. We extracted DNA from microdissected tissue obtained from 200 surgically resected NSCLCs. *KDR* single nucleotide polymorphism (SNP) 889G/A (rs2305948), SNP 1416A/T (rs1870377) and SNP -37A/G (rs2219471) were genotyped by PCR-based sequencing. *KDR* and *VEGF* copy number were examined by quantitative (q)-PCR. Protein expressions of VEGF and VEGFR2, and CD34 for microvascular density (MVD) analysis were studied by immunohistochemistry. SNP genotypes, genes copy number and protein expression were correlated with NSCLC clinico-pathological features, including overall survival (OS) and recurrence-free survival (RFS).
Results. *KDR* 1416 AT/TT genotypes had significantly improved OS (HR = 0.56,

95% CI 0.33 to 0.96, $P = 0.035$) compared with *KDR* 1416 AA in all NSCLC patients in the multivariate analysis with adjustment of histology and tumor stage. In lung adenocarcinomas, *KDR* 1416 AT/TT genotypes and *KDR* -37AG/GG were associated with a favorable OS (HR = 0.43, 95% CI 0.20 to 0.92, $P = 0.029$; HR = 0.47, 95% CI 0.23 to 0.96, $P = 0.039$, respectively). Strikingly, *KDR* -37 AG/GG genotypes predicted a superior OS compared with *KDR* -37 AA genotype in the NSCLC patients treated with adjuvant therapy (HR = 2.45, 95% CI 1.06 to 5.66, $P = 0.036$). No genotype in *KDR* SNPs was associated with RFS in NSCLC patients. Gene copy gain of *KDR* and *VEGF* were detected respectively in 34/91 (37.4%) and 2/91 (2.2%) NSCLC tumors. Gene copy gain of *KDR* was associated with a poor overall survival in NSCLC patients (HR= 2.96, 95% CI 1.41 to 6.24, $P = 0.004$). Furthermore, tumors with gene copy gain of *KDR* showed significantly higher cytoplasmic ($P = 0.013$) and membrane ($P = 0.009$) VEGFR2 protein expression, lower cytoplasmic VEGF expression ($P = 0.044$), and higher MVD ($P = 0.018$) and larger vessel areas ($P = 0.033$) compared with tumors lacking *KDR* gene copy gain.

Conclusion. Our findings suggest an association between *KDR* SNP genotypes and survival in NSCLC patients receiving adjuvant therapy. *KDR* copy number gain was frequently identified in NSCLC and was associated with worse survival in NSCLC patients. (Supported by grant US DoD W81XWH-07-1-0306).

American Association for Cancer Research
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Expression of Stem Cell Markers in Non-Small Cell Lung Carcinoma (NSCLC) and Correlation with Clinico-pathologic Features

Ping Yuan¹, Carmen Behrens¹, Jiaoti Huang², Monica Spinola³, Ludmila Prudkin¹, Wenli Dong¹, Guosheng Yin¹, Cesar Moran¹, Edward Kim¹, Bin-Bing Zhou⁴, John Minna³, Ignacio Wistuba¹.

University of Texas, M. D. Anderson Cancer Center, Houston TX¹, University of California Los Angeles², University of Texas Southwestern Medical Center Dallas TX³, and Wyeth Pharmaceuticals, Pearl River, NY⁴.

Background. Cancer stem cells (CSCs) or cancer-initiating cells represent a minor population of self-renewing tumor cells which are believed to play an important role in tumor development, metastasis and resistance to therapy. Although some CSC markers have been described in NSCLC, there is no comprehensive characterization of multiple CSC markers in this disease. Our aim was to investigate the patterns of protein expression of a panel of CSC-related markers in a large series of NSCLCs, and correlate those findings with patients' clinico-pathologic characteristics.

Methods. We examined protein expression by immunohistochemistry (IHC) of 287 NSCLCs (178 adenocarcinoma, and 109 squamous cell carcinomas (SCC)) of a panel of nine CSC markers: EZH2, SOX2, CD24, CD44, C-kit, HEY1, Shh, BMI-1 and Oct3/4. The patterns of expression of these markers were correlated with patients and tumors' clinico-pathologic characteristics, including disease outcome. In the adenocarcinoma, CSC markers expression was correlated with the *EGFR* and *KRAS* mutation status of the tumors.

Results. EZH2, SOX2, CD44, CD24 and C-kit were detected in the tumor cells. The pattern of Immunohistochemistry staining varied in each markers, EZH2 and SOX2 in nuclei, CD44 and CD24 in membrane, and C-kit in cytoplasm. Rare cases had HEY1 and Shh expression. We didn't find BMI-1 and Oct3/4 expression in our cases. The expression for these markers correlated with certain clinico-pathologic characteristics, including tumor histology, pathological stage, and patients' smoking history. EZH2, SOX2 and CD44 expressions were significantly higher in SCC than adenocarcinoma ($P<0.001$), and CD24 expressions was significantly higher in adenocarcinoma than in SCC ($P<0.05$). Patients with tobacco history showed significantly higher EZH2, SOX2, and CD24 expression compared with patients without tobacco history ($P<0.001$, $P<0.05$). The presence of *EGFR* mutation in lung adenocarcinoma correlated significantly with low EZH2 ($P=0.03$) and high CD44 ($P=0.032$) expression. We identified a subset of NSCLCs having membrane CD44+/CD24 low or negative expression. Interestingly, in multivariate analysis using the expression scores as a continuous variable, high nuclear expression of EZH2 correlated significantly with worse recurrence free survival (HR=1.004; $P=0.021$) and overall survival (HR=1.006; $P=0.017$) in patients with stage I/II adenocarcinoma who didn't

receive pretreatment. Patients with CD44+/CD24- expression had worse overall survival than other subtype in squamous cell carcinoma in male (HR=2.935, p=0.037)

Conclusions. We provide a characterization of multiple CSC markers in a large series of NSCLCs. Our findings indicate that a different pattern of CSC markers expression is detected in adenocarcinoma and squamous cell carcinomas of the lung, and their expression correlates with patients' clinico-pathologic features, including survival. The understanding of the role of CSCs in NSCLC tumor development and progression may provide opportunities to design novel strategies to prevent and treat this disease.

Supported by grants DoD W81XWH-04-1-0142, W81XWH-07-1-0306, and UT-Lung SPORE P50CA70907.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 5166

Presentation
Title: **Sex determining region Y-box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung**

Presentation
Time: Wednesday, Apr 21, 2010, 8:00 AM -11:00 AM

Location: Exhibit Hall A-C, Poster Section 12

Poster
Section: 12

Poster
Board
Number: 25

Author
Block: Ping Yuan¹, Humam Kadara¹, Carmen Behrens¹, Ximing Tang¹, Denise Woods¹, Luisa M. Solis¹, Jiaoti Huang², Monica Spinola³, Wenli Dong¹, Guosheng Yin¹, Junya Fujimoto¹, Edward Kim¹, Yang Xie³, Luc Girard³, Cesar Moran¹, Waun Ki Hong¹, John D. Minna³, Ignacio Ivan Wistuba¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²David Geffen School of Medicine, Los Angeles, CA; ³The University of Texas Southwestern Medical Center, Dallas, TX

Abstract
Body: Non-small cell lung cancer (NSCLC) represents the majority (85%) of lung cancers and is comprised mainly of adenocarcinomas and squamous cell carcinomas (SCCs). The sequential pathogenesis of lung adenocarcinomas and SCCs occurs through dissimilar phases as the former tumors typically arise in the lung periphery whereas the latter normally arise near the central airway. We assessed the expression of SOX2, an embryonic stem cell transcriptional factor that also plays important roles in the proliferation of basal tracheal cells and whose expression is restricted to the main and central airways and bronchioles of the developing and adult mouse lung, in NSCLC by various methodologies. Here, we found that *SOX2* mRNA levels, from various published datasets, were significantly elevated in lung SCCs compared to adenocarcinomas (all $p < 0.001$). Moreover, a previously characterized *OCT4/SOX2/NANOG* signature effectively separated lung SCCs from adenocarcinomas following integration with two independent publicly available gene expression microarray datasets and which correlated with increased *SOX2* mRNA in SCCs. Immunohistochemical analysis of various histological lung tissue specimens demonstrated marked nuclear SOX2 protein expression in all normal bronchial epithelia, alveolar bronchiolization structures and premalignant lesions in SCC development (hyperplasia, dysplasia and carcinoma *in situ*) and absence of expression in all normal alveoli and atypical adenomatous hyperplasias. Moreover, SOX2 protein expression was greatly higher in lung SCCs compared to adenocarcinomas following analyses in two independent large tissue microarray (TMA) sets (TMA set I, $n=287$; TMA set II, $n=511$ both $p < 0.001$). Furthermore, amplification of *SOX2* DNA was detected in 20% of lung SCCs tested ($n=40$) and in none of the adenocarcinomas ($n=17$). Our findings highlight a cell-lineage gene expression pattern for the stem cell transcriptional factor SOX2 in the pathogenesis of lung SCCs and raise the intriguing possibility of the growth addiction of lung SCCs specifically to SOX2-dependent pathways. Supported in part by grants from the Department of Defense (W81XWH-04-1-0142 and W81XWH-07-1-03060), and the Specialized Program of Research Excellence in Lung Cancer grant P50CA70907.

Running title: SOX2 abnormalities in NSCLC

American Association for Cancer Research
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Appendix: PROSPECT Database Screenshots (new development shown on exhibits 11-14.)

1) Clinical module: Patient Information, Social History, Medical History

Admin

Projects

Histo-Pathology Lab

Logout

MRN

Name

Ben%

Benjamin G Sample

Path#

☒ Patient Completed

662526

Status: Ready...

Other Malignancy

Treatment

Staging

Follow up

Pathology

All Clinical TO EXCEL

Patient Information

EXPORT TO EXCEL

Last

Sample

Middle

G

First

Benjamin

MDAH

662526

Gender

Male

Race

White

DOB

12/31/1936

City

Montgomery

State

TX

Enter Date

06/22/2006

Zip Code

77356

Country

USA

Age

Last Visit Date

Death or Alive

Alive

Lost To Follow up

n/a

Chemotherapy

n/a

Radiotherapy

n/a

Recurrence

n/a

Social History

EXPORT TO EXCEL

Smoking History

☐ Yes ☐ No

Pack Years

0

Are you currently smoking?

☐ Yes ☒ No

If no, date quit smoking

Age Quit Smoking

Age started smoking regularly

18

Overall Smoking Years

Average number of cigarettes smoked per day?

20

Actual Smoking Years

If there was a quit-smoking period, total time during the smoking years.

0 yrs., 0 mos.

Pack Years

Asbestos Exposure

☐ Yes ☐ No

Alcohol History

☒ Yes ☐ No

No. of Drinks / Month

0

Medical History

EXPORT TO EXCEL

Hypertension

☐ Yes ☒ No

Heart Problem

☐ Yes ☒ No

Thyroid

☐ Yes ☐ No

Diabetes

☐ Yes ☒ No

Renal Insufficiency

☐ Yes ☐ No

Asthma

☐ Yes ☒ No

DVT

☐ Yes ☒ No

Pulmonary Embolism

☐ Yes ☒ No

Mild hemoptysis

☐ Yes ☐ No

Radiation Fibrosis

☐ Yes ☐ No

Hepatic Problem

☐ Yes ☐ No

COPD




☐ Yes ☐ No

Other

Hypertension, Hyperlipidemia, Depression, Gastric Reflux

Date COPD Dx

2) Clinical module: Other Malignancy

OtherMalignancy: 499975				Status: Ready...	
Click here to add <input type="text" value="1"/> more row(s).				Save it	Cancel it
				Open it	Save and Close
EXPORT TO EXCEL					
Malig. ID	Patient ID	Dx Date	Malig. Detail		Treatment
 45	113	1/1/1900	Period: Synchronous to lung cancer (within a year) Organ: Lung Histology: Adenocarcinoma Comments: NED: <input type="checkbox"/> NED Date: 1/1/1900	Surgery: <input type="checkbox"/> Date: 3/12/2003 Chemo: <input type="checkbox"/> Date: 1/1/1900 Radio: <input type="checkbox"/> Date: 1/1/1900	
 46	113	1/1/1900	Period: n/a Organ: Skin Histology: Melanoma Comments: NED: <input type="checkbox"/> NED Date: 1/1/1900	Surgery: <input type="checkbox"/> Date: 4/20/2004 Chemo: <input type="checkbox"/> Date: 1/1/1900 Radio: <input type="checkbox"/> Date: 1/1/1900	
 49	113	1/1/1900	Period: After lung cancer Organ: Skin Histology: Squamous Cell Carcinoma Comments: NED: <input type="checkbox"/> NED Date: 1/1/1900	Surgery: <input checked="" type="checkbox"/> Date: 2/9/2005 Chemo: <input type="checkbox"/> Date: 1/1/1900 Radio: <input type="checkbox"/> Date: 1/1/1900	

3) Clinical module: Treatment: Surgery, Chemotherapy, Radiotherapy and Other Treatments

Status: Ready...

Surgery: 523912

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Surgery ID	Patient ID	Surgery Date	Is MDA	Surgery Procedure	Comments	Margin Left
29	142	3/24/2003	<input type="checkbox"/>	n/a		<input type="checkbox"/>

Chemotherapy

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Chemo ID	Chemo Type	Chemo Date	Drug	Tumor Size	Response	Comments
29	Is MDA: <input type="checkbox"/> Chemo. Type: Neoadjuvant	Start: 1/1/1900 Stop: 1/1/1900	A: Carboplatin B: Paclitaxel C: n/a #crs: 0	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	
143	Is MDA: <input type="checkbox"/> Chemo. Type: 1st Line	Start: 7/29/2006 Stop: 9/16/2003	A: Cisplatin B: Etoposide C: n/a #crs: 0	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	

Radiotherapy

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)


Radio ID	Treatment Option	Radio Date	Tumor Size	Response	Comments
37	Is MDA: <input type="checkbox"/> Site1: L-lung Site2: n/a Site3: n/a Treatment Option: 1st Line	Start: 7/29/2003 Stop: 9/16/2003	Before(CT) 0 After(CT) 0 % Reduction NaN Before(Patho) 0 After(Patho) 0 % Reduction NaN	Clin.: n/a CT: n/a Patho: n/a	

Other Treatment



[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Treatment ID	Patient ID	Surgery Date	Other Treatment	Comments
8	142	5/5/2004	Craniotomy	

4) Clinical module: Staging

Staging:				Status: Ready...									
Click here to add <input type="text" value="1"/> more row(s).				Save it		Cancel it		Open it		Save and Close		Excel Report	
Stage ID	Patient ID	Staging Date	Current Situation				Clin. T	Clin. N	Clin. M	Clin. Stage	Pleu Eff	Malig Pl Eff	
 1	44	1/1/1900	Multiple primary tumor different histology				T2	N0	M1	IB	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Pathology													
Tumor Specimens													
Patient ID	Accession	Surgical Date	Single Wedge	Multiple Wedge	Single Segmentectomy	Multiple Segmentectomy	Lobectomy	Bilobectomy	Pneumonectomy	# Nodules	Tum ID		
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	23		
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	26		
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	27		
<div> <input type="text"/> <input type="button" value="Previous"/> <input type="button" value="Next"/> </div>													
Dx Specimens													
Patient ID	AccessionNo	Path Type	Event	Dx Specimen Date	Specimen Type	Tumor Site	Specimen Avail						
44	SB-1111	Primary	Dx Specimen	01/19/2007	CORE BIOPSY	LLL							
Metastasis Specimens													
Patient ID	AccessionNo	Path Type	Event	Met Date	Specimen Type	Tumor Site	Specimen Avail						

5) Clinical module: Follow up

Follow up: 127855				Status: Ready...	
Click here to add <input type="text" value="1"/> more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL					
Person Review <input type="text" value="n/a"/>				Date Updated <input type="text"/>	
Fu ID	Patient ID	Fu Date	Fu Detail		
			Status: <input type="text" value="recurrence"/>		
			If "no change or no recurrence": Form of contact: <input type="text"/>		
			If "recurrence": Form of contact: <input type="text" value="visit to clinic"/> Date of recurr: <input type="text" value="1/1/1900"/>		
			<input type="text" value="Thoracic Lymph"/>		
			Site of Recur: <input type="text" value="n/a"/> If Lung: <input type="text" value="n/a"/>		
			<input type="text" value="n/a"/>		
			Recur Biopsy: <input type="checkbox"/> Image Date: <input type="text" value="1/1/1900"/>		
			If "death": Death Date: <input type="text"/> Cause of Death: <input type="text"/>		
			Info by: <input type="text"/>		
			Comments: <input type="text"/>		
			Status: <input type="text" value="no change"/>		
			If "no change or no recurrence": Form of contact: <input type="text" value="visit to clinic"/>		
			If "recurrence": Form of contact: <input type="text"/> Date of recurr: <input type="text"/>		
			<input type="text"/>		
			Site of Recur: <input type="text"/> If Lung: <input type="text"/>		
			<input type="text"/>		
			Recur Biopsy: <input type="checkbox"/> Image Date: <input type="text"/>		
			If "death": Death Date: <input type="text"/> Cause of Death: <input type="text"/>		
			Info by: <input type="text"/>		
			Comments: <input type="text"/>		
	422	55	2/4/2004		
	298	55	1/4/2006		

6) Pathological module: Tissue Pathological Data

- Primary Diagnosis Specimen
- Primary Surgical Specimen
- Metastasis Diagnosis Specimen
- Metastasis Surgical Specimen

Admin
Projects
Histo-Pathology Lab
Logout

Select a participant
Status: Ready...

Hist Dx TO EXCEL
All Pathology TO EXCEL

[Click here](#) to add 1 more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

				Accession	Biopsy	Event
				563 SB-1111	Primary	Surgical Specimen
				564 SB-2222	Primary	Dx Specimen
				566 SS-1112	Metastasis	Surgical Specimen
				567 S-04-23495	Primary	Surgical Specimen
				568 SB-01-1234	Metastasis	Dx Specimen

Dx Specimen

[Click here](#) to add 1 more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Dx Specimen ID	Obtained Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail

Surgical Specimen

[Click here](#) to add 1 more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Surg ID	Acc. ID	Accession No	Surgical Date	Single Wedge	Multi Wedge	Single Segmen tectomy	Multi Segmen tectomy	Lobec tomy	Bilo bectomy	Pneumo nectomy	No of Nodules	T	N	M
			6	567	S-04-23495	5/11/2005								2	n/e	n/e	n/e

Met Specimen

[Click here](#) to add 1 more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Met Specimen ID	Met Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail
			6	1/1/1900	566	SS-1112	n/a	n/a	

7) Pathological module: Histology

Histology		Status: Ready	
<div>UpdateDelete</div>			
Hist ID:	5	Dx Specimen ID:	22
Met Specimen ID:	0	Tumor ID:	0
Histology Dx:	Metastasis to lung		
Metastasis (to the lung) Dx:	Carcinoma	Other Tumoral Characteristics (Hist Dx)	
Carcinoma Met Site:	Breast	Necrosis %	23
		Fibrosis %	3
		Inflammation	Severe

8) Pathological module: Staging and Tumor Information

Tumor: s-04-23495 (6)				Status: Ready...																																																															
Click here to add <input type="text" value="1"/> more row(s). Save it Cancel it Open it Save and Close Excel Report																																																																			
		Tumor ID	Surgical ID	Specimen Type	Tumor Site	Localization																																																													
		DX		<input type="text" value="23"/>	<input type="text" value="6"/>	<input type="text" value="n/a"/>	<input type="text" value="n/a"/>																																																												
		DX		<input type="text" value="26"/>	<input type="text" value="6"/>	<input type="text" value="Bilobectomy"/>	<input type="text" value="n/a"/>																																																												
		DX		<input type="text" value="27"/>	<input type="text" value="6"/>	<input type="text" value="Lobectomy"/>	<input type="text" value="n/a"/>																																																												
Type of Tissue Available <input type="checkbox"/> Bronchus Extrapulmonary <input type="checkbox"/> Principal <input type="checkbox"/> Lobar Intrapulmonary <input type="checkbox"/> with Cartilage <input type="checkbox"/> w/o Cartilage <input type="checkbox"/> Bronchiole <input type="checkbox"/> Alveoli		Tumor Size(cm) <input type="text" value="0"/> <input type="checkbox"/> Tumor Invasion <input type="checkbox"/> Pleural <input type="checkbox"/> Neural <input type="checkbox"/> Vascular <input type="checkbox"/> Other <input type="checkbox"/> Margin Positive <input type="checkbox"/> Bronchial <input type="checkbox"/> Parenchymal <input type="checkbox"/> Soft Tissue <input type="checkbox"/> Other		Pathological T Pathological T: <input type="text" value="n/a"/> <input type="checkbox"/> Tx T1: <= 3 cm. T2: > 3 cm or <= 3cm and/or attached to visceral pleura <input type="checkbox"/> Pleural Attached T3: <input type="checkbox"/> Parietal Pleura <input type="checkbox"/> Mediastinal Pleura <input type="checkbox"/> Chest Wall <input type="checkbox"/> Mediastinal Fat <input type="checkbox"/> Pericardium <input type="checkbox"/> Phrenic Nerve <input type="checkbox"/> Vagus Nerve <input type="checkbox"/> Sympathetic Chain <input type="checkbox"/> Atelectasis Entire Lung T4: <input type="checkbox"/> >1 Nodal <input type="checkbox"/> Invades to Great Vessel <input type="checkbox"/> Invades to Heart <input type="checkbox"/> Invades to Trachea <input type="checkbox"/> Invades to Carina <input type="checkbox"/> Invades to Esophagus <input type="checkbox"/> Invades to Vertebral Bones <input type="checkbox"/> Associated with Malignant Pleural Effusion <input type="checkbox"/> Satellite Tumor Nodule (size)																																																															
Field Study Normal bronchial epithelium and premalignant lesions <input type="checkbox"/> Normal Bronchial Epithelium <input type="checkbox"/> Normal Lung Parenchyma <input type="checkbox"/> Hyperplastic Alveoli <input type="checkbox"/> Bronchial Hyperplasia <input type="checkbox"/> Squamous Metaplasia <input type="checkbox"/> Mild Squamous Dysplasia <input type="checkbox"/> Moderately Squamous Dysplasia <input type="checkbox"/> High Squamous Dysplasia <input type="checkbox"/> In Situ Squamous Carcinoma <input type="checkbox"/> Atypical Adenomatous Hyperplasia <input type="checkbox"/> Tumorlet				Lymph Node Metastasis <input type="checkbox"/> Lymph Node Involvement Analysis of Lymph Node Station <table border="1"> <thead> <tr> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>S1</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S2</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S3</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S4</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> </tr> <tr> <td>S5</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S6</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S7</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S8</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> </tr> <tr> <td>S9</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S10</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S11</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S12</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> </tr> <tr> <td>S13</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>S14</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>NS</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> <td>Total</td> <td><input type="text" value="0"/></td> <td><input type="text" value="0"/></td> </tr> </tbody> </table> <input type="checkbox"/> Positive Contralateral LN mtt Path N: <input type="text" value="n/a"/> Type of LN metastasis Size of Metastasis <input type="checkbox"/> Intranodal <input type="checkbox"/> Capsular Min. Size: <input type="text" value="0"/> Max. Size: <input type="text" value="0"/> <input type="checkbox"/> Subcapsular <input type="checkbox"/> Perinodal					+	Total		+	Total		+	Total		+	Total	S1	<input type="text" value="0"/>	<input type="text" value="0"/>	S2	<input type="text" value="0"/>	<input type="text" value="0"/>	S3	<input type="text" value="0"/>	<input type="text" value="0"/>	S4	<input type="text" value="0"/>	<input type="text" value="0"/>	S5	<input type="text" value="0"/>	<input type="text" value="0"/>	S6	<input type="text" value="0"/>	<input type="text" value="0"/>	S7	<input type="text" value="0"/>	<input type="text" value="0"/>	S8	<input type="text" value="0"/>	<input type="text" value="0"/>	S9	<input type="text" value="0"/>	<input type="text" value="0"/>	S10	<input type="text" value="0"/>	<input type="text" value="0"/>	S11	<input type="text" value="0"/>	<input type="text" value="0"/>	S12	<input type="text" value="0"/>	<input type="text" value="0"/>	S13	<input type="text" value="0"/>	<input type="text" value="0"/>	S14	<input type="text" value="0"/>	<input type="text" value="0"/>	NS	<input type="text" value="0"/>	<input type="text" value="0"/>	Total	<input type="text" value="0"/>	<input type="text" value="0"/>
	+	Total		+	Total		+	Total		+	Total																																																								
S1	<input type="text" value="0"/>	<input type="text" value="0"/>	S2	<input type="text" value="0"/>	<input type="text" value="0"/>	S3	<input type="text" value="0"/>	<input type="text" value="0"/>	S4	<input type="text" value="0"/>	<input type="text" value="0"/>																																																								
S5	<input type="text" value="0"/>	<input type="text" value="0"/>	S6	<input type="text" value="0"/>	<input type="text" value="0"/>	S7	<input type="text" value="0"/>	<input type="text" value="0"/>	S8	<input type="text" value="0"/>	<input type="text" value="0"/>																																																								
S9	<input type="text" value="0"/>	<input type="text" value="0"/>	S10	<input type="text" value="0"/>	<input type="text" value="0"/>	S11	<input type="text" value="0"/>	<input type="text" value="0"/>	S12	<input type="text" value="0"/>	<input type="text" value="0"/>																																																								
S13	<input type="text" value="0"/>	<input type="text" value="0"/>	S14	<input type="text" value="0"/>	<input type="text" value="0"/>	NS	<input type="text" value="0"/>	<input type="text" value="0"/>	Total	<input type="text" value="0"/>	<input type="text" value="0"/>																																																								

9) Pathological module: Tissue Bank (Frozen and Paraffin)

Tissue Bank: s-04-23495				Status: Ready...			
Click here to add <input type="text" value="1"/> more row(s). Save it Cancel it Open it Save and Close Excel Report							
		Frozen	FFPE	TBID	Collection Date	Frozen Avail	FFPE Avail
				<input type="text" value="6"/>	<input type="text" value="4/12/2008"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
				<input type="text" value="8"/>	<input type="text" value="4/13/2008"/>	<input type="checkbox"/>	<input type="checkbox"/>
				<input type="text" value="9"/>	<input type="text" value="4/12/2008"/>	<input type="checkbox"/>	<input type="checkbox"/>
Frozen							
Click here to add <input type="text" value="1"/> more row(s). Save it Cancel it Open it Save and Close Excel Report							
		ID		Tissue	Blood	Pleural	
		Frozen ID: <input type="text" value="0"/> TB ID: <input type="text" value="6"/> SPORE No: <input type="text" value="0"/> TID No: <input type="text" value="0"/>	Normal Lung	<input type="checkbox"/>	DNA Conc <input type="text" value="0"/>	DNA Conc <input type="text" value="0"/>	DNA Conc <input type="text" value="0"/>
			Tumor	<input type="checkbox"/>	DNA Vol <input type="text" value="0"/>	DNA Vol <input type="text" value="0"/>	DNA Vol <input type="text" value="0"/>
			Bronchus	<input type="checkbox"/>	DNA Quality <input type="text" value="n/a"/>	DNA Quality <input type="text" value="n/a"/>	DNA Quality <input type="text" value="n/a"/>
			LN	<input type="checkbox"/>	RNA Conc <input type="text" value="0"/>	RNA Conc <input type="text" value="0"/>	RNA Conc <input type="text" value="0"/>
			Serum	<input type="checkbox"/>	RNA Vol <input type="text" value="0"/>	RNA Vol <input type="text" value="0"/>	RNA Vol <input type="text" value="0"/>
			Lymphocyte	<input type="checkbox"/>	RNA Quality <input type="text" value="n/a"/>	RNA Quality <input type="text" value="n/a"/>	RNA Quality <input type="text" value="n/a"/>
			Pleural	<input type="checkbox"/>	Prot Conc <input type="text" value="0"/>	Prot Conc <input type="text" value="0"/>	Prot Conc <input type="text" value="0"/>
					Prot Vol <input type="text" value="0"/>	Prot Vol <input type="text" value="0"/>	Prot Vol <input type="text" value="0"/>
					Prot Quality <input type="text" value="n/a"/>	Prot Quality <input type="text" value="n/a"/>	Prot Quality <input type="text" value="n/a"/>
FFPE							
Click here to add <input type="text" value="1"/> more row(s). Save it Cancel it Open it Save and Close Excel Report							
		FFPE ID	TB ID	Cabinet	Tray	Block	Slide
		<input type="text" value="0"/>	<input type="text" value="6"/>	<input type="text"/>	<input type="text"/>	<input type="text" value="0"/>	<input type="text" value="0"/>

10) Dictionary module

Admin
Projects
Histo-Pathology Lab
Logout

Dictionaries

Country
Update

Page 1 [2] [3] [4] [5] [...]

	Dictionaries	Set Order
+		
	Afghanistan	1
	Albania	2
	Algeria	3
	American Samoa	4
	Andorra	5
	Angola	6
	Anguilla	7
	Antarctica	8
	Antigua and Barbuda	9
	Arctic Ocean	10
	Argentina	11
	Armenia	12
	Aruba	13
	Ashmore and Cartier Islands	14
	Atlantic Ocean	15
	Australia	16
	Austria	17
	Azerbaijan	18
	Bahamas	19

Page 1 [2] [3] [4] [5] [...]

Page 1 of 14

11) Query Tool

Admin
Projects
Histo-Pathology Lab
Logout

Query Generator

Query Template Management

select query:
Select
open
enter query name:
private

Run
Save
Overwrite
Delete
Rename

Status: Ready

Query Criteria
Reset
Create a new query
Go!

Field ID	Field Name	Operator	Field Value	And/Or
+	0	Select...		and

Field Selection
Reset

Selected Fields	Unselect	Clear	Ordering Fields	Unselect	Clear	Sorting Fields	Unselect	Clear

Clear
Reset
Select All

☒ All
☐ Accession
☐ Chemotherapy
☐ FFPE
☐ Frozen Bank
☐ General Info
☐ Histological Dx
☐ Other Malignancy
☐ Other Treatment
☐ Patient
☐ Radiology
☐ Staging
☐ Surgery
☐ Tissue Bank
☐ Tumor

<input type="checkbox"/> Acc No	<input type="checkbox"/> Date COPD Dx	<input type="checkbox"/> Invasion Heart	<input type="checkbox"/> Middle Name	<input type="checkbox"/> Protocol ID	<input type="checkbox"/> Site
<input type="checkbox"/> Accession No	<input type="checkbox"/> Date Other Tx	<input type="checkbox"/> Invasion Trachea	<input type="checkbox"/> Mild Hemoptysis	<input type="checkbox"/> Race	<input type="checkbox"/> Slide
<input type="checkbox"/> AccessionID	<input type="checkbox"/> Date Quit Smoke	<input type="checkbox"/> Invasion Vertical Bones	<input type="checkbox"/> Mild Squamous Dysplasia	<input type="checkbox"/> Radiation Fibrosis	<input type="checkbox"/> Smoked >= 100
<input type="checkbox"/> Actuar	<input type="checkbox"/> Date Screening	<input type="checkbox"/> Invasion Vessel	<input type="checkbox"/> Mod Squamous Dysplasia	<input type="checkbox"/> Radio Clin Response	<input type="checkbox"/> Both Tissue Margin
<input type="checkbox"/> Address1	<input type="checkbox"/> Date Staging	<input type="checkbox"/> Is Chemo	<input type="checkbox"/> More Than 1 Nodal	<input type="checkbox"/> Radio Comments	<input type="checkbox"/> Solid
<input type="checkbox"/> Address2	<input type="checkbox"/> Date Surgery	<input type="checkbox"/> Is NED	<input type="checkbox"/> Multiple Segmentectomy	<input type="checkbox"/> Radio CT Response	<input type="checkbox"/> Specimen Avail
<input type="checkbox"/> Age Start Smoke	<input type="checkbox"/> Diabetes	<input type="checkbox"/> Is Radio	<input type="checkbox"/> Multiple Wedge	<input type="checkbox"/> Radio MDA	<input type="checkbox"/> Specimen Type
<input type="checkbox"/> Alveoli	<input type="checkbox"/> DOB	<input type="checkbox"/> Is Surgery	<input type="checkbox"/> N Clin	<input type="checkbox"/> Radio Patho Response	<input type="checkbox"/> SPORL_No
<input type="checkbox"/> Asbestos	<input type="checkbox"/> Dx	<input type="checkbox"/> Last Name	<input type="checkbox"/> Necrosis	<input type="checkbox"/> Radio Site	<input type="checkbox"/> Squamous Metaplasia
<input type="checkbox"/> Asthma	<input type="checkbox"/> Dx Date	<input type="checkbox"/> LN	<input type="checkbox"/> NED Date	<input type="checkbox"/> Radio Start Date	<input type="checkbox"/> State
<input type="checkbox"/> Asthmatic Lung	<input type="checkbox"/> Dx Specimen Date	<input type="checkbox"/> LN Involve	<input type="checkbox"/> No Drinks a Mo	<input type="checkbox"/> Radio Stop Date	<input type="checkbox"/> Study ID
<input type="checkbox"/> Atypical Adeno Hyper	<input type="checkbox"/> Dx Variant	<input type="checkbox"/> LN Max Size	<input type="checkbox"/> No of Nodules	<input type="checkbox"/> Radio Treatment Option	<input type="checkbox"/> Sub Type
<input type="checkbox"/> Avg Cig a Day	<input type="checkbox"/> DxP Actuar	<input type="checkbox"/> LN Met	<input type="checkbox"/> No Smoke Mo	<input type="checkbox"/> Radio Tumor Size After CT	<input type="checkbox"/> Surgery
<input type="checkbox"/> BAC	<input type="checkbox"/> DxP BAC	<input type="checkbox"/> LN Met Capsular	<input type="checkbox"/> No Smoke Years	<input type="checkbox"/> Radio Tumor Size After Patho	<input type="checkbox"/> Surgery Comments
<input type="checkbox"/> Bilelectomy	<input type="checkbox"/> DxP Carcinoma Met Site	<input type="checkbox"/> LN Met Intramural	<input type="checkbox"/> Normal Bronch Ipi	<input type="checkbox"/> Radio Tumor Size Before Patho	<input type="checkbox"/> Surgery Procedures
<input type="checkbox"/> Black	<input type="checkbox"/> DxP Dx	<input type="checkbox"/> LN Met Perimural	<input type="checkbox"/> Normal Lung Parenchyma	<input type="checkbox"/> Radiotherapy	<input type="checkbox"/> Surgical Date
<input type="checkbox"/> Blood DNA Conc	<input type="checkbox"/> DxP Dx Variant	<input type="checkbox"/> LN Met Subcapsular	<input type="checkbox"/> Normal Lung	<input type="checkbox"/> Radio Tumor Size Before CT	<input type="checkbox"/> Symp Chain
<input type="checkbox"/> Blood DNA Quality	<input type="checkbox"/> DxP Fibrosis	<input type="checkbox"/> LN Met Size	<input type="checkbox"/> NS =	<input type="checkbox"/> Renal Insuff	<input type="checkbox"/> T Clin
<input type="checkbox"/> Blood DNA Volume	<input type="checkbox"/> DxP Grade	<input type="checkbox"/> Lobectomy	<input type="checkbox"/> NS Analyzed	<input type="checkbox"/> S1	<input type="checkbox"/> TID
<input type="checkbox"/> Blood Protein Conc	<input type="checkbox"/> DxP Inflammation	<input type="checkbox"/> Lymphocyte	<input type="checkbox"/> NSCLC Type	<input type="checkbox"/> S1 -	<input type="checkbox"/> Thyroid
<input type="checkbox"/> Blood Protein Quality	<input type="checkbox"/> DxP Met Dx	<input type="checkbox"/> M Clin	<input type="checkbox"/> On Study	<input type="checkbox"/> S10	<input type="checkbox"/> TID_No
<input type="checkbox"/> Blood Protein Volume	<input type="checkbox"/> DxP Microcapillary	<input type="checkbox"/> Malign PI Eff	<input type="checkbox"/> Other Margin	<input type="checkbox"/> S10 -	<input type="checkbox"/> Tumor Line
<input type="checkbox"/> Blood RNA Conc	<input type="checkbox"/> DxP Necrosis	<input type="checkbox"/> Malignant Pleural Eff	<input type="checkbox"/> Other MDA Rx	<input type="checkbox"/> S11	<input type="checkbox"/> Tissue DNA Conc
<input type="checkbox"/> Blood RNA Quality	<input type="checkbox"/> DxP NSCLC Type	<input type="checkbox"/> Margin =	<input type="checkbox"/> Other Met History	<input type="checkbox"/> S11 -	<input type="checkbox"/> Tissue DNA Quality
<input type="checkbox"/> Blood RNA Volume	<input type="checkbox"/> DxP Papillary	<input type="checkbox"/> Margins Residual	<input type="checkbox"/> Other Treatment Comments	<input type="checkbox"/> S12	<input type="checkbox"/> Tissue DNA Volume
<input type="checkbox"/> Bronch Hyperplasia	<input type="checkbox"/> DxP Sarcoma Met Site	<input type="checkbox"/> MDA	<input type="checkbox"/> Papillary	<input type="checkbox"/> S12 -	<input type="checkbox"/> Tissue Protein Conc
<input type="checkbox"/> Bronchial Margins	<input type="checkbox"/> DxP Solid	<input type="checkbox"/> MDAH	<input type="checkbox"/> Parenchymal Margins	<input type="checkbox"/> S13	<input type="checkbox"/> Tissue Protein Quality
<input type="checkbox"/> Bronchiol	<input type="checkbox"/> DxP Sub Type	<input type="checkbox"/> Mediastinal Pleura	<input type="checkbox"/> Pleural Pleura	<input type="checkbox"/> S13 -	<input type="checkbox"/> Tissue Protein Volume
<input type="checkbox"/> Bronchus	<input type="checkbox"/> Eligible	<input type="checkbox"/> Mediastinal Fat	<input type="checkbox"/> Path M	<input type="checkbox"/> S14	<input type="checkbox"/> Tissue RNA Conc
<input type="checkbox"/> Bronchus Sample	<input type="checkbox"/> Endobron	<input type="checkbox"/> Met Date	<input type="checkbox"/> Path N	<input type="checkbox"/> S14 -	<input type="checkbox"/> Tissue RNA Quality
<input type="checkbox"/> Cancers	<input type="checkbox"/> Event	<input type="checkbox"/> Met Dx	<input type="checkbox"/> Path T	<input type="checkbox"/> S2	<input type="checkbox"/> Tissue RNA Volume
<input type="checkbox"/> Carcinoma Met Site	<input type="checkbox"/> Ever Consumed Alcohol	<input type="checkbox"/> Met Specimen Avail	<input type="checkbox"/> Path Type	<input type="checkbox"/> S2 -	<input type="checkbox"/> Total -
<input type="checkbox"/> Cardio	<input type="checkbox"/> Expiratory Lobar	<input type="checkbox"/> Met Specimen Type	<input type="checkbox"/> Patho N	<input type="checkbox"/> S3	<input type="checkbox"/> Total Analyzed
<input type="checkbox"/> Cardio DVT	<input type="checkbox"/> Expiratory Principal	<input type="checkbox"/> Met Tumor Size	<input type="checkbox"/> Patho T	<input type="checkbox"/> S3 -	<input type="checkbox"/> Tumor
<input type="checkbox"/> Cdx Phone	<input type="checkbox"/> FFPE	<input type="checkbox"/> MetP Actuar	<input type="checkbox"/> Patient Comments	<input type="checkbox"/> S4	<input type="checkbox"/> Tumor
<input type="checkbox"/> Chemo Drug A	<input type="checkbox"/> Fibrosis	<input type="checkbox"/> MetP BAC	<input type="checkbox"/> Patient ID	<input type="checkbox"/> S4 -	<input type="checkbox"/> Tumor Invasion
<input type="checkbox"/> Chemo Drug B	<input type="checkbox"/> First Name	<input type="checkbox"/> MetP Carcinoma Met Site	<input type="checkbox"/> Pericardium	<input type="checkbox"/> S5	<input type="checkbox"/> Tumor Invasion Neural
<input type="checkbox"/> Chemo Drug C	<input type="checkbox"/> Frozen	<input type="checkbox"/> MetP Dx	<input type="checkbox"/> Phrenic Nerve	<input type="checkbox"/> S5 -	<input type="checkbox"/> Tumor Invasion Other
<input type="checkbox"/> Chemo MDA	<input type="checkbox"/> Gender	<input type="checkbox"/> MetP Dx Variant	<input type="checkbox"/> Pleural	<input type="checkbox"/> S6	<input type="checkbox"/> Tumor Invasion Pleural
<input type="checkbox"/> Chemo Start Date	<input type="checkbox"/> Grade	<input type="checkbox"/> MetP Fibrosis	<input type="checkbox"/> Pleural Attached	<input type="checkbox"/> S6 -	<input type="checkbox"/> Tumor Invasion Vascular
<input type="checkbox"/> Chemo Stop Date	<input type="checkbox"/> Hepatic	<input type="checkbox"/> MetP Grade	<input type="checkbox"/> Pleural DNA Conc	<input type="checkbox"/> S7	<input type="checkbox"/> Tumor Localization
<input type="checkbox"/> Chemo Type	<input type="checkbox"/> High Squamous Dysplasia	<input type="checkbox"/> MetP Inflammation	<input type="checkbox"/> Pleural DNA Quality	<input type="checkbox"/> S7 -	<input type="checkbox"/> Tumor Organ Site
<input type="checkbox"/> Chemotherapy	<input type="checkbox"/> Histology	<input type="checkbox"/> MetP Met Dx	<input type="checkbox"/> Pleural DNA Volume	<input type="checkbox"/> S8	<input type="checkbox"/> Tumor Site
<input type="checkbox"/> Chest Wall	<input type="checkbox"/> Home Phone	<input type="checkbox"/> MetP Microcapillary	<input type="checkbox"/> Pleural Eff	<input type="checkbox"/> S8 -	<input type="checkbox"/> Tumor Size
<input type="checkbox"/> City	<input type="checkbox"/> Hyperplastic Alveolar	<input type="checkbox"/> MetP Necrosis	<input type="checkbox"/> Pleural Protein Conc	<input type="checkbox"/> S9	<input type="checkbox"/> Tumor Type
<input type="checkbox"/> Clin Stage	<input type="checkbox"/> In Situ Squamous Carcinoma	<input type="checkbox"/> MetP NSCLC Type	<input type="checkbox"/> Pleural Protein Quality	<input type="checkbox"/> S9 -	<input type="checkbox"/> Tumorlet
<input type="checkbox"/> Collection Date	<input type="checkbox"/> Inflammation	<input type="checkbox"/> MetP Papillary	<input type="checkbox"/> Pleural Protein Volume	<input type="checkbox"/> Sarcoma Met Site	<input type="checkbox"/> Tx
<input type="checkbox"/> COPD	<input type="checkbox"/> Intra v Carriage	<input type="checkbox"/> MetP Sarcoma Met Site	<input type="checkbox"/> Pleural RNA Conc	<input type="checkbox"/> Sarcoma Tumor	<input type="checkbox"/> Vagus Nerve
<input type="checkbox"/> Country	<input type="checkbox"/> Intra v Carriage	<input type="checkbox"/> MetP Solid	<input type="checkbox"/> Pleural RNA Quality	<input type="checkbox"/> Serum	<input type="checkbox"/> Work Phone
<input type="checkbox"/> Current Situation	<input type="checkbox"/> Invasion Carcin	<input type="checkbox"/> MetP Sub Type	<input type="checkbox"/> Pleural RNA Volume	<input type="checkbox"/> Single Segmentectomy	<input type="checkbox"/> Zip
<input type="checkbox"/> Current Smoker	<input type="checkbox"/> Invasion Esophagus	<input type="checkbox"/> Microcapillary	<input type="checkbox"/> Pneumectomy	<input type="checkbox"/> Single Wedge	

12) The query results page

Query Name: Adenocarcinoma																										
Record Count: 13																										
Accession No	Biobectomy Date	Surgical Specimen Event Date	Lobectomy	Met Date	Met Specimen Avail	Met Specimen Type	Met Specimen Site	Multiple Segmentectomy	Multiple Wedge	No. of Nodes	Path M	Path N	Path I	Path Type	Pneumonec	Single Segmentectomy	Single Wedge	Specimen Avail	Specimen Type	Surgical Date	Tumor Site	Acinar	BAG	Carcinoma Met Site	Dx	C
S-03-000995	False	Surgical Specimen	False					False	False	1	M0	N0	T2	Primary	False	False	True			1/8/2003 12:00:00 AM	95	5	n/a	NSCLC		
S-03-015760	False	Surgical Specimen	False					False	False	1	M0	N0	T1	Primary	False	False	True			4/2/2003 12:00:00 AM	80	0	n/a	NSCLC		
S-03-018420	False	Surgical Specimen	True					False	False	1	M0	N2	T1	Primary	False	False	False			4/15/2003 12:00:00 AM	0	0	n/a	NSCLC	n	
S-03-029423	False	Surgical Specimen	False					False	False	1	M0	N2	T2	Primary	False	False	True			4/29/2003 12:00:00 AM	60	0	n/a	NSCLC		
S-03-024985	False	Surgical Specimen	True					False	False	1	M0	N0	T1	Primary	False	False	False			5/23/2003 12:00:00 AM	20	80	n/a	NSCLC		
S-03-052181	False	Surgical Specimen	True					False	False	3	M0	N2	T4	Primary	False	False	False			10/23/2003 12:00:00 AM	0	0	n/a	NSCLC	n	
S-04-051605	False	Surgical Specimen	True					False	False	1	M0	N0	T2	Primary	False	False	False			10/6/2004 12:00:00 AM	25	75	n/a	NSCLC		
S-04-056678	False	Surgical Specimen	True					False	False	2	M1	N0	T2	Primary	False	False	True			11/3/2004 12:00:00 AM	35	50	n/a	NSCLC		
S-05-020799	False	Surgical Specimen	True					False	False	1	M0	N0	T2	Primary	False	False	False			4/19/2005 12:00:00 AM	0	0	n/a	NSCLC		
S-05-029775	False	Surgical Specimen	True					False	False	1	M0	N1	T3	Primary	False	False	False			6/6/2005 12:00:00 AM	0	0	n/a	NSCLC		
S-05-050475	False	Surgical Specimen	True					False	False	1	M0	N2	T2	Primary	False	False	False			9/19/2005 12:00:00 AM	0	0	n/a	NSCLC	n	
S-05-065030	False	Surgical Specimen	False					False	False	1	M0	N0	T2	Primary	False	False	True			12/7/2005 12:00:00 AM	0	0	n/a	NSCLC		
S-05-068150	False	Surgical Specimen	True					False	False	1	M0	N1	T2	Primary	False	False	False			12/27/2005 12:00:00 AM	0	0	n/a	NSCLC	c	

13) The example of the Excel reports.

	K	L	M	N	O	P	Q
1	Specimen Avail	Surgical Date	Single Wedge	Multiple Wedge	Single Segmentectomy	Multiple Segmentectomy	Lobectomy
2							
3							
4							
5		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
6		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
7		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
8		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
9							
10							
11							
12							
13		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
14		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
15	FALSE						
16	TRUE						
17							
18							
19		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
20							
21							
22							
23							
24							
25							
26		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
27							
28							
29							
30							
31							
32							
33							
34							
35							

14) Patient Summary Report

